Reliability of Fatigue Measures in an Overhead Work Task: A Study of Shoulder Muscle Electromyography and Perceived Discomfort

Kristopher M. R. Hager

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Master of Science
In
Industrial and Systems Engineering

Dr. Maury A. Nussbaum, Chair

Dr. Kari L. Babski-Reeves

Dr. Tonya L. Smith-Jackson

October 31, 2003

Blacksburg, VA

Keywords: discomfort, shoulder, fatigue, electromyography, reliability

Reliability of Fatigue Measures in an Overhead Work Task: A Study of Shoulder Muscle Electromyography and Perceived Discomfort

Kristopher M. R. Hager

(ABSTRACT)

This study was conducted to measure the reliability of fatigue measures in an intermittent overhead work task. Fatigue measures included several EMG based parameters and subjective discomfort ratings through use of the Borg CR-10 scale. This study was part of a larger existing study that simulates overhead work in an automobile manufacturing plant. Ten participants used a drill tool to perform an overhead tapping task for one hour at a height relative to individual anthropometry.

Reliability indexes, including Intraclass Correlation Coefficients, Standard Errors of Measurement, and Coefficients of Variation were determined for each fatigue measure for each of three shoulder muscles (anterior deltoid, middle deltoid, and trapezius). High reliability implies repeatable results, and precise and credible methods. Conversely, measurement error and subject variability can lead to low reliability of measures.

The results indicated that ratings of perceived discomfort (RPD) parameters (slope and final rating) showed relatively high reliability. Intercepts for mean power frequency (MnPF), median power frequency (MdPF), and root means square (RMS) also showed very high reliability. Actual slopes for MnPF, MdPF, and RMS showed low reliability overall, and normalizing slopes did not necessarily improve reliability. Taking the absolute value of slopes led to a noticeable increase in reliability. RPD slope did not correlate with any of the EMG slopes.

The high reliability of RPD parameters allows for its inexpensive application to the industrial setting for similar overhead tasks. The reliability of EMG intercepts implies consistent methods; however the reliability of overall EMG trends is suspect if the slope is not reliable. Some EMG slope parameters show promise; however, more research is needed to determine if these parameters are reliable for complex tasks.

ACKNOWLEDGMENTS

I would firstly like to thank Dr. Nussbaum. He is one heck of an advisor, who has always guided me in the right direction and who always had his door open to my questions and concerns. I would also like to thank my committee members, Dr. Babski-Reeves and Dr. Smith-Jackson. Their questions and comments helped add quality to this work.

I am grateful to my parents and brother, who have all helped get me to where I am today and where I'll be going tomorrow.

Special thanks to Yassierli, Dadi, Deepti, Kesh, and Yen, who have all provided either intellectual or moral support. Thanks to the whole VT HFES gang, all of whom added to my graduate experience, making it more enjoyable and just a great place to learn.

Finally, I would like to acknowledge the hidden wonders of Virginia Tech and Blacksburg, which has been a home away from home.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Work-Related Musculoskeletal Disorders Overview

Injuries that affect the biomechanical structure and integrity of human body tissues, such as nerves, tendons, ligaments, and muscles, are often referred to as Musculoskeletal Disorders (MSDs). MSDs that are developed or worsened in the work place are commonly known as Work-Related Musculoskeletal Disorders (WMSDs). Examples of WMSDs include back strain, shoulder tendonitis, and carpal tunnel syndrome. WMSDs can cause pain or numbness, temporary or permanent disability, lost work time, and an increase in workers' compensation costs (NIOSH, 1997).

1.1.1 WSMD Prevalence

Current injury statistics indicate the high prevalence of WMSDs. In 2001, the Bureau of Labor Statistics (BLS) reported over 5.2 million nonfatal occupational injuries and illnesses, of which 2.6 million cases involved lost work days. Of the 5.2 million cases, over 4.9 million were injuries, and over 333,000 were illness cases which involved repeated trauma such as carpal tunnel syndrome. In 1996, the Occupational Safety and Health Administration (OSHA) estimated that workers compensation for WMSDs resulted in approximately $20 billion in direct costs and about $80 billion more in indirect costs. In effect, the large numbers of WMSDs and their cost indicate the necessity for intervention.

1.1.2 Shoulder Injuries

Shoulder injuries are an area of particular concern in the workplace. The BLS (2001) reported over 88,000 occupational injuries and illnesses involving the shoulder that caused days away from work. Of these, over 55,000 were MSD related and were second in number of occurrences to the back. After the wrist and abdomen, injuries to the shoulder caused the longest time away from work with a median of 12 days (BLS, 2001). The large number of shoulder injuries in the workplace and the result of these injuries indicate a need for research in this area.

1.1.3 Approaches to the WSMD Problem

Private industry, the U.S. government, and universities implement programs, research results, and regulations to improve the health and safety of workers. Many such programs involve ergonomics, which can be defined as science of designing the task, tools, and work environment to fit the capabilities of the worker. Specifically, occupational biomechanics applies ergonomic principles to improve worker performance and decrease risk of WMSDs.

Biomechanics is the study of forces acting on the human body and the effects of these forces on the body’s tissues, fluids, or materials (Radwin et al., 2002). Figure 1 provides a conceptual model of factors that may play a role in development of WMSDs. A worker may experience biomechanical loading which can exceed internal tolerances and have various consequences. Of importance in this research is the relationship between muscle fatigue and its effect on pain and discomfort which may lead to WMSD development.

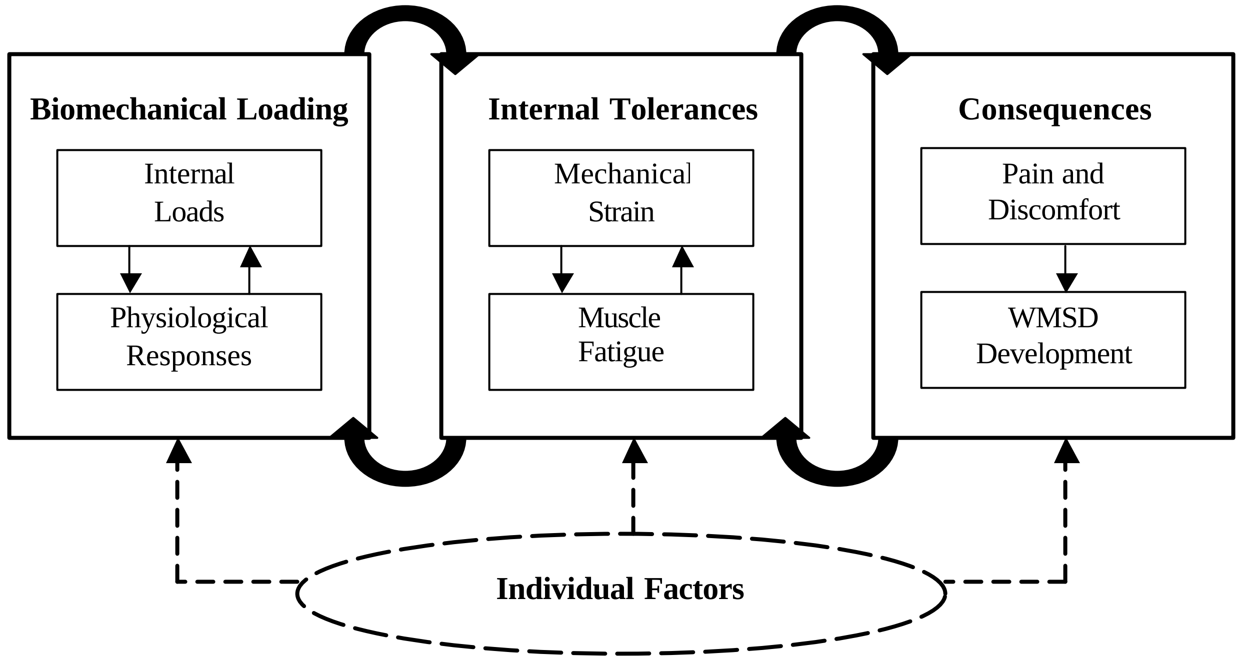


Figure 1: Conceptual model of factors that may play a role in the development of WMSDs. (Modified from National Research Council and the Institute of Medicine, 2001; Radwin et al., 2002)

1.2 Muscle Fatigue Research

1.2.1 Fatigue and WSMD risk Relationship

Muscle fatigue, defined when the muscle fails to maintain the required or expected force (Dimitrova, 2003; Green, 1996), is a broadly researched topic with a variety of motivations. For example, Hui et al. (2001) conducted a study on nurses in a geriatric ward, with a stated motivation that a given load on the back muscles in combination with back muscle fatigue can contribute to back strain. In a study of fatigue resulting from overhead work, Nussbaum (2001) assumed that muscle fatigue can be used as an indicator of injury risk. Thus, by minimizing fatigue, the risk of injury may be reduced.

The relationship between fatigue and injury risk, however, has not been clearly defined. It has been suggested that the association of task demands and worker capacity may influence the occurrence of fatigue, discomfort, and injury (Dempsey, 1998). This connection may give merit to fatigue being an indicator of risk for injury. It might further be assumed that tasks leading to fatigue are associated with an increased risk of musculoskeletal injury, and that fatigue is a way to measure the physical demands of a given task. If this is a valid assumption, then fatigue research has been a major factor in helping to reduce the large number of yearly occupational injuries.

Another motivation to study fatigue is its effect on functional capacity, which decreases due to prolonged tissue loads or frequent exertions (Chaffin et al. 1999). Fatigue is also associated with a decline in muscle force resulting from metabolic changes within the muscle and impaired activation (Chaffin et al. 1999). For a variety of Manual Material Handling (MMH) tasks, wherein a necessary muscle force output is required, moderate to high levels of fatigue will have a negative impact on performance of any given pushing, pulling, lifting, carrying, or holding task. An example of this is conveyed in a study of lifting dynamics, in which Chen (1999) observed that lifting strategies differed significantly when participants were subjected to upper extremity fatigue. Chen (1999) also showed that some fatigued participants were at increased risk of injury to the lower back.

1.2.2 Measures of Fatigue

There are several commonly used methods in fatigue research. Objective measures include use of electromyography (EMG), maximum strength, and endurance times. Subjective measures include use of ratings scales such as Borg's 10-point scale (Borg, 1982).

In order to have accurate and precise results, fatigue measures must be valid and reliable. Several studies have been completed in order to assess the validity of various fatigue measures such as EMG and Borg Scales (Gerdle et al., 2000; Larsson et al., 1999; Shen et al., 1997).

Other work, such as by Rainoldi et al. (1999) and Falla et al. (2002), has indicated that there are few reliability studies of muscle fatigue measures with appropriate statistical support. The work here focuses on reliability of several common fatigue measures.

1.3 Overview of Reliability

1.3.1 Defining Reliability

Reliability of an experiment implies that it has repeatable results. The Merriam-Webster Dictionary (2002) defines reliability as "the extent to which an experiment, test, or measuring procedure yields the same results on repeated trials." Reliability also involves the precision of methods and the level of credibility placed on results (Aarass et al., 1996). Reliability is important, because issues such as measurement error and subject variability can have a negative impact on statistical results and interpretation of these results.

Reliability should not to be confused with validity. A study or method is valid if it measures what it is designed to measure. A method can be reliable, but not valid, but in order for a method to be valid it must be reliable (Fagarasanu et al., 2002). Therefore, the reliability of a method will affect its level of validity.

There are several methods to assess reliability. These include, split-halves method, alternate-form method, internal consistency, inter-rater method, and test-retest method. This present research focuses on test-retest reliability, which measures consistency over time. An example of test-retest reliability involves the participant taking the same test during two different points in time. This form of reliability is considered more costly than the others, but it is a simple and clear reliability method, and can be widely applied to diverse research areas in biomechanics, specifically fatigue research.

1.3.2 Indexes for Measuring Reliability

There are a number of indexes for reliability and there is much conflicting literature on which index is most appropriate for use. A review of recent research showed that Pearson's correlation coefficient (*r*), though commonly used in the past studies, is not an acceptable form of measuring test-retest reliability (Denegar et al., 1993; Keating, 1998). Pearson's *r* can overestimate reliability, because the systematic bias in the repeated measures will go undetected. Pearson's *r* measures the strength of the relationship between test and retest measures, and not the agreement between them (Bland et al., 1986; Larssona et al., 2003). A more appropriate measure than Pearson's *r* would be intraclass correlation coefficients (ICC).

ICC is the ratio of the between-subjects variance divided by the total variance (Denegar et al., 1993). Shrout and Fleiss (1979) discuss six forms of ICC and how to apply them. These forms are labeled (1,1), (2,1), (3,1), (1,K), (2,K), and (3,K). Test-retest reliability with a single rater and a 2-way ANOVA is estimated in Equation 1. ICC can range from 0 to 1, wherein 0 indicates no reliability and 1 indicates perfect reliability. A negative ICC indicates that the within-subject variance exceeded the between-subjects variance and is equivalent to an ICC of 0, or no reliability.

$$\mathrm{ICC}(2,1) = \frac{(\mathrm{BMS} - \mathrm{EMS})}{\mathrm{BMS} + (k-1)\mathrm{EMS} + k[(\mathrm{TMS} - \mathrm{EMS})/N]} \qquad (1)$$

BMS = between-subjects mean square
EMS = error mean square
TMS = trial mean square
k = number of trials or evaluators
n = number of subjects

Another commonly used reliability index is the standard error of measurement (SEM), which estimates the precision of what is being measured (Denegar et al., 1993), and is defined as the standard deviation times the square root of one minus the ICC (Equation 2). It can be considered an "average error" that is presented in the same units as what is being measured (Keating, 1998). A high SEM value in relation to the measurements indicates a high level of error, and thus, the test and retest values are not reproducible.

$$\mathrm{SEM} = s\sqrt{1-r} \tag{2}$$

s = standard deviation of the measurements
r = ICC

The recent trend in reliability research is to report ICC and SEM together. ICC is a measure of relative reliability, while SEM is a measure of absolute reliability. The combination is useful, because in some instances ICC can produce misleadingly high levels of reliability. ICC takes into account the between subject variance and will provide a high reliability if there is a large variance between subjects (Keating, 1998). In these cases, SEM will yield a relatively high value and help the experimenter to recognize that the test-retest values actually have low reproducibility.

A few studies report the coefficient of variation (CV), defined as the SEM over the mean value (Equation 3). The CV allows for comparisons between different variables and methods (Elfving et al., 1999) by scaling the SEM by the mean. A low CV indicates high precision, while a high CV indicates low precision.

$$\mathrm{CV} = \frac{\mathrm{SEM}}{\mathrm{mean}} \times 100 \tag{3}$$

1.3.3 Classifying Reliability Indexes

Classification of the level of reliability, represented by the ICC index, is often an area of disagreement among different researchers. Researchers disagree on ranges of reliability and associated descriptions ("poor," "good," "excellent," etc.) of these levels. Table 1 presents several commonly used interpretations of ICC levels. Fleiss (1986) has suggested that there is really no universal method for categorizing ICC levels, given that errors in determining reliability may occur even with high ICC levels.

Table 1: Classification of ICC level of reliability from psychology, medical, and ergonomic related research.

Source	Interpretation
Bartko et al. (1966)	0-.6, poor, .6-.8 good, .8-1.0 excellent
Landis et al. (1977)	0-.2 slight, .21-.4 fair, .41-.6 moderate, .61-.8 substantial, .81-1.0 almost perfect
Fleiss (1986)	0-.4 poor, .4-.75 fair to good, .75-1.0 excellent
Sleivert et al. (1994)	0-.59, poor, .6-.79 fair, .8-1.0 good
Shrout (1998)	0-.1 virtually none, .11-.4 slight, .41-.6 fair, .61-.8 moderate, .81-1.0 substantial
Stokdijk et al. (2000)	0-.39 poor, .4-.59 fair, .6-.74 good, .75-1.0 excellent
Koumantakis et al. (2002)	0-.69, poor, .7-.79 fair, .8-.89 good, .9-.99 high

From the classifications in Table 1, the following general guidelines seem warranted: 0-.39 poor, .40-.59 fair, .60-.79 good, .8-1.0 excellent (negatives fall under poor). This system is mainly a combination of Bartko et al. (1966), and Stokdijk et al. (2000). The system used by Stokdijk et al. (2000) seemed appealing, because one can use the terms "good" and "excellent" to indicate relatively high reliability. "Fair" indicates that though the reliability may not be high, it is at a level wherein improvement of methods or larger sample sizes may increase reliability. "Poor" indicates that reliability is so low that the parameter in check would not be very useful or applicable. The system is then modified using the concept by Bartko et al. (1966) that "excellent" reliability is a little more stringent (.8 to 1.0 rather than .75 to 1.0), and an adjustment is made to "good reliability as well (upper bound of .74 increases to .79).

For the remainder of this document, reliability from existing studies will be classified using the scheme Table 1. While not claimed as a perfect classification, this scheme allows for consistency and reading ease since most studies vary in classifications or don't use a classification system. In addition, if the SEM and CV reported is high enough to question a relatively high ICC, this factor will be mentioned.

1.3.4 Applications of Reliability

Test-retest reliability research can be applied to a variety of areas such as medicine, psychology, and biomechanics. Reliability is used to evaluate instruments, such as a grading system for glenohumeral anthropathy (Ilg et al., 2001), or to evaluate methods for assessing rotator cuff strength (Hayes et al., 2002). It is used in a clinical setting wherein comparisons are made of chronic low pack pain patients with normal subjects (Lariviere et al., 2000). Reliability also helps to evaluate the reproducibility of EMG and subjective ratings while performing isometric trunk extensions (Elfving et al., 1999). Of particular interest in this research is the reliability of EMG parameters and subjective discomfort during a fatiguing task.

1.4 Electromyography Reliability

In biomechanics studies, EMG has three main applications, which include detection of muscle activation, formation of a muscle force-EMG relationship, and use as an index of fatigue (De Luca, 1997). This research focuses on the latter of the three applications. Muscle fatigue is often associated with a decrease in median and mean power frequencies (MnPF, MdPF) and/or an increase in root mean square (RMS) levels (Rainoldi, 1999). These changes are typically found during simple (static) exertions, though less consistent results have been found for more complex (intermittent and/or dynamic) activities.

1.4.1 Issues Affecting EMG Reliability

Inconsistent methods used for EMG acquisition often have a negative impact on EMG reliability. Time-of-day bias, laboratory temperature, and prior fatigue (Koumantakis et al., 2000) can affect EMG parameters. It is also well known that source and amplifier input impedance, types of electrodes used, electrode contact area (Jukka et al., 1975), and interelectrode distance (Elfving et al., 2002) affect reliability as well. Understanding and accounting for these issues will help reduce the effect of experimental methods on reliability of EMG.

1.4.2 EMG Parameters

Most existing reliability studies reveal that initial MnPF and MdPF values tend to be highly reliable while their slopes have poor reliability (Peach et al., 1998; Elfving et al., 1999; Koumantakis et al., 2001). Some studies, though, have found that MdPF and MnPF slopes have good reliability (Dedering et al., 2000; Falla et al., 2002) These mixed results indicate the biomechanics applicability of the initial values for MnPF and MdPF, but raise questions as to the acceptability of use of the rate of decline for detecting fatigue.

There exist fewer reliability studies of EMG RMS than MnPF and MdPF. One such study (Larsson et al., 2003), measuring the EMG reliability of dynamic maximum concentric knee extensions, indicated that RMS and MnPF parameters were comparably reliable. Other studies showed that the reliability of initial MnPF, MdPF, RMS values were comparable, though MnPF and MdPF slopes tended to be more reliable than RMS slope (Kollmitzer et al., 1999; Lavierre et al., 2002). Nargol et al. (1999) suggested that RMS slopes may be less reliable, because they are more load dependent than MnPF and MdPF. On the contrary, a study by Koumantakis et al. (2001), which involved maintaining isometric contractions of the paraspinal muscles at 40, 50, and 60% MVE, indicated that the MdPF slope was more reliable than RMS slope at 50% MVE, but less reliable at 60%. More research is clearly needed to compare the reliability of MnPF, MdPF, and RMS parameters.

Two other types of parameters that have received limited study are normalized slope and absolute value of slope. Lariviere (2002) suggests that normalizing EMG slope parameters to intercepts (or initial values) can account inter-individual differences such as subcutaneous tissue thickness which may affect the consistency of EMG readings. Therefore, normalizing slopes may increase the reliability of the EMG slope parameter. In a study by Luttman et al. (1996), trapezius EMG was measured from surgeons performing endoscopy, and decreases in MdPF were associated with fatigue (as expected), but increases in MdPF were associated with force increases. Force increases can indicate either recovery or fatigue, as more force is required to achieve a similar output. In some cases, MdPF might then show positive and negative slope (increase or decrease) as a result of fatigue. Based on the equation for ICC and CV, taking the absolute of slopes may prevent low reliability values due to differences in negative and positive slopes.

1.5 Borg Scale Reliability

Since 1962, Borg has been developing scales to measure perceived exertion. The Borg 10-point Category Ratio (CR-10) scale (Borg, 1982) is a one dimensional scale from 0 to 10 with verbal anchors. Compared to previous Borg scales, the CR-10 is used to measure a wider range of psychophysical intensities. Specifically, the CR-10 scale has been widely used to obtain subjective provide ratings of perceived discomfort (RPD). The Borg scale was used for this study over other subjective discomfort scales, since it is a well known and widely used scale (Borg, 1982; Kumar et al., 1999).

1.5.1 Issues Affecting Borg Scale Reliability

Ambiguity, inconsistent methods, time effects, and motivation are some many sources of poor reliability for subjective ratings. Shen et al. (1997) reported that while using the Borg CR-

10 scale to measure discomfort, some participants complained about the ambiguity of the Borg scale wording, such as “somewhat,” and at times used the scale inconsistently. Ambiguity can be caused in part by inconsistent methods, which include inconsistency in explanation and demonstration of the Borg scale. Time effects occur when a participant is asked to give a subjective rating after having performed some tasks, and the participant has to rely on memory (Annett, 2002). Motivation can also affect subjective ratings, which may cause participants to under or overestimate their actual rating. Motivation can be partially influenced by time of day effects, wherein some people may have higher motivation in either the morning or night.

To help prevent poor reliability, methods should be consistent. The Borg scale needs to be explained clearly and consistently across participants. The experiments should be conducted on the same time each day, or as close to the same time on each day as possible. Additionally, there must be little to no delay in a participant providing the subjective rating after performing the task.

1.5.2 Borg Scale Reliability Studies

Although reliability studies of the Borg scale to assess perceived effort has been determined, there are very limited studies on the use of this scale for discomfort. Shen et al. (1997) reported that Borg CR-10 had only fair reliability in reporting seated pressure discomfort. In a study measuring dyspnea during exercise on a cycle ergometer (Mador et al., 1995), participant ratings of degree of discomfort evoked by breathing had a CV of 9% at maximal workload and a CV of 13.6% at 66% of the maximum workload, which indicates good reliability.

Studies that utilize Borg scale intensities to assess perceptions other than discomfort seem to indicate high levels of reliability. An experiment requiring participants to maintain an isometric trunk extension found excellent reliability for subjective ratings of fatigue and exertion (Elfving et al., 1999). Wos et al. (1988) found excellent levels of reliability for subjective ratings of short duration hand-arm vibrations. Although both these studies show excellent reliability, more research is needed measuring the intensity of discomfort.

Some studies have indicated Borg subjective ratings to assess discomfort have been correlated to other objective measures. Gorman et al. (1999) showed that when used to rate breathing discomfort, Borg scores were linearly related to progressive increasing end-tidal CO_2 over time. A study by Wikstrom (1992) indicated a good correlation between Borg scale ratings for discomfort and both whole-body vibration level and degree of body twist (angle) when driving in various conditions. Dedering et al. (1999) demonstrated that Borg ratings of perceived fatigue was correlated with MnPF and MdPF slope in a static back extension task. There is still, however, a lack of studies which assess the correlation between Borg scale ratings for discomfort and EMG fatigue parameters during more realistic dynamic tasks.

1.6 Time Span Between Test and Retest

For most reliability studies, test and retests are either conducted within-day (on the same day) or between-day (on a different day). Time between tests in within-day experiments has ranged from 3-120 minutes (Kollmitzer et al., 1999; Rainoldi et al., 2001; Ebenbichler et al., 2002), and morning to afternoon (Elfving et al., 1999). Time between tests for between-day experiments has ranged from one to two weeks (Mannion et al., 1994), five to seven days (Dedering et al., 2000), and two days to two weeks (Lariviere, 2002).

Longer time intervals have noticeable effects on reliability. In a study of isometric knee extension tests, Kollmitzer et al. (1999) found that an interval of 6 weeks between tests had lower EMG reliability than 90min and 3min intervals. Time intervals of greater than 3 months seem to significantly reduce reliability (Koumantakis et al., 2002). Most studies report that within-day experiments, using EMG based measures of fatigue, have good to excellent reliability while between-day experiments vary between excellent and poor reliability (Elving et al., 1999; Ebenbichler, 2002).

There is a definite need for more reliability studies of longer test-retest intervals. Certain training programs that involve fatiguing workers in an industrial setting or patients in a clinical setting require intervals of days or weeks to evaluate these programs (Dedering et al., 2000). One study (Rainoldi, et al., 1999) had participants perform isometric contractions with rest intervals of 5 minutes, and found that some participants needed longer rest between tests. Experiments that contain longer fatiguing conditions may require longer rests periods to prevent residual fatigue effects and cannot be repeated on the same day or even the following day. As a result, the present study explored between-days reliability.

1.7 Automotive Research Study

For this research, an automotive task with methods similar to that described in Sood et al. (2002) was selected. This task is an example of one posing a potential risk of WMSD in an industrial setting. The purpose of this fatigue study was to determine the effects of three different overhead work heights (low, middle, and high) on shoulder fatigue. Participants performed a dynamic intermittent task simulating automotive assembly work. Throughout the trials, EMG was recorded from three shoulder muscles (anterior deltoid, middle deltoid, and trapezius), and discomfort was rated using Borg's CR-10.

For test-retest evaluations, the middle height was used. Participants had their retest session two days to two weeks following their initial test session, and the reliability of several fatigue measures was determined.

CHAPTER 2. RESEARCH OBJECTIVES

2.1 Rationale for the Study

WMSDs, and specifically shoulder injuries, are highly prevalent in the workplace. Muscle fatigue research is one of many ergonomic methods seeking to reduce the occurrence of WMSDs. All research, fatigue included, must use reliable measures.

Reliability is important, because measurement error and subject variability are directly related to the reliability of measures. Low reliability of measures may negatively affect the validity of measures (Fagarasanu et al., 2002). If reliability is not accounted for, statistical results can be misinterpreted.

A test-retest reliability study was conducted in the context of a simulated overhead task. There are several reliability studies involving the shoulder, but there are limited reliability studies for shoulder fatigue measures specifically. There is also limited work on the reliability of measures obtained during dynamic tasks (Larsson, 1999), as most studies focus on isometric contractions (Rainoldi et al., 1999). This study involved a dynamic intermittent overhead task, and the results were intended to add to existing shoulder and dynamic reliability studies.

This study analyzed the reliability of EMG parameters and ratings using the Borg CR-10 scale. Current studies provide conflicting evidence as to the reliability of EMG parameters (Kollmitzer et al., 1999; Larsson et al., 2003). There are also limited studies analyzing the reliability of Borg CR-10 scale for discomfort (Shen et al., 1997). This study sought to add to current data on reliability of both fatigue measures.

The results of this work determined if EMG and RPDs are reliable fatigue measures for research involving dynamic or shoulder related tasks. These measures are applicable to research and industry if results indicate good to excellent levels of reliability and a low SEM/CV. On the contrary, poor to fair levels of reliability and a moderate to high SEM implies that different measures of fatigue may be necessary for repeatable results. Additionally, the CV can help determine which measure is the most reliable.

Based on SEM values and the t-distribution, 95% confidence interval was reported for each fatigue measure (Equation 4; for v = 9 degrees of freedom, t = 2.262). Variation within this interval can be considered normal variability with 95% confidence. A worker with an RPD higher than the upper limit of this confidence interval may be at risk for WSMD development, or in the case of a new worker, may not have yet experienced work hardening. It should be noted, however, that only relatively low SEM values can be used in this manner, as higher values are more likely to indicate measurement error (Denegar et al., 1993).

$$C.I. = \pm t_{.975,v} * SEM \quad (4)$$

2.2 Research Question

EMG and discomfort ratings were collected for three shoulder muscles (anterior deltoid, medial deltoid, and trapezius) during a relatively complex task (intermittent and dynamic). The associated research questions were:

-What is the level of reliability of several EMG based measures?

-What is the level of reliability for the Borg CR-10 scale for discomfort measures?

-What is the level of correlation between Borg slope and various EMG slopes?

CHAPTER 3. EXPERIMENTAL METHODS

3.1 Experimental Design

This study was part of an existing larger study designed to assess the relationship between overhead working height and associated shoulder fatigue. In the larger study, three heights, determined relative to each participant's anthropometry, were studied. The heights were selected based on observations and measurements made at an automobile assembly plant (Figure 2), and subsequent analyses of videotape recordings of several automobile assembly tasks. For the current study, participants performed only the middle height task on two different sessions.

Data from the last four participants (2 females, 2 males) from the larger study were used in the current study. Though part of the larger study, these participants were selected with the same screening procedures as the current study (described below in Section 3.2). Rew et al. (2000) discusses several possible limitations of secondary data analysis, which includes not having full access to data, documentation of data, the investigator not being involved in the design of the original study, and the data reflecting only the original investigators' questions and perspectives. None of these limitations apply here, because the investigator for the current study was also involved in the previous study and had full access to the original data and understood its applications.

The two test and retest sessions were separated by at least 48 hours to ensure that there was no residual fatigue from the previous session, which might otherwise confound data obtained in a subsequent session. The two sessions were separated by no longer than one week as a longer time span may affect reliability. To ensure against time of day bias, both sessions were conducted at approximately the same time (morning, afternoon, or evening).

Two anthropometric measures were taken for the dominant hand/arm to determine the subject-specific task height. These measures were: hand height when the upper arm is held horizontal and the elbow is flexed at 90^{o} (H1), and the hand height with arm in full extension (overhead reach) with shoulders parallel to ground (H2). Hand height was measured to the center of the grip. An illustration of hand heights is provided in Figure 3. The middle height is the sum of H1 and 40% of the difference between H1 and H2 (H1 + 0.40[H1-H2]).

Figure 2: Representative overhead work in an automobile assembly plant.

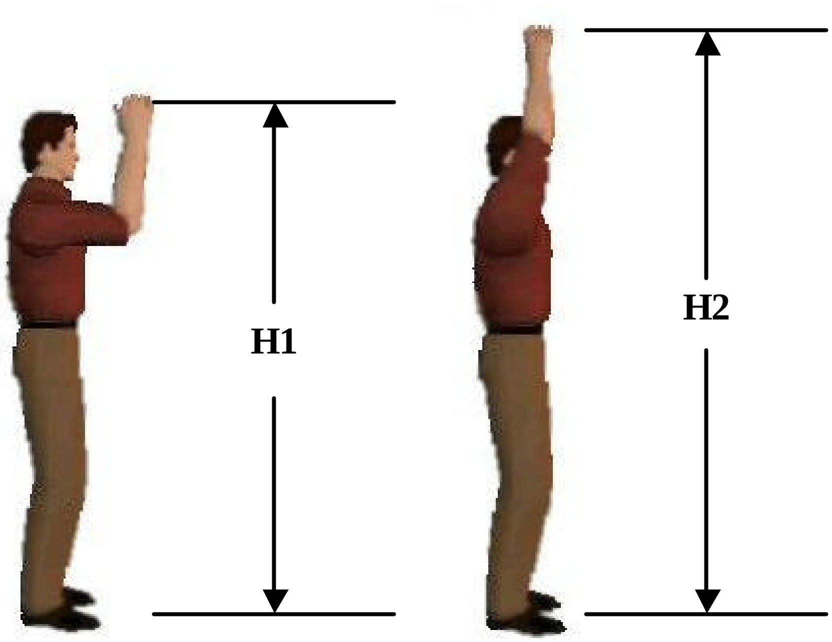


Figure 3: Hand heights at H1 and H2.

3.2 Participants

Reliability studies suggest a wide range of sample sizes. Several studies recognize that reliability studies can be time consuming and expensive (Charter, 1999; Perkins et al., 2000). Two conflicting studies, emphasizing precise measurement of error, suggest that the sample size would range from 8 to 200 (Hopkins, 2000) or require a minimum of 400 subjects (Charter, 1999). Donner et al. (1987) suggested that to have an 80% (ß = .2) certainty of achieving a significant result (getting ICC = .8, while the null hypothesis is ICC = 0) at the a = .05level with two sessions, at least 8 participants are required. To achieve significant results by getting an ICC of less than .8 under the same conditions (a = .05, ß = .2, while null hypothesis is ICC of 0), however, requires more retest sessions and participants (Donner et al., 1987). Given the large sample size requirements from these studies a different approach was used.

Some reliability studies agree that sample size should be based on a power of 80% to detect a meaningful difference in group means with an a error set at .05 (Donner et al., 1987; Cichetti, 1999). With a = .05, ß = .2, and k = 1.5 standard deviations (where k is the number of number of standard deviations between means to be detected with high probability), a power approach indicates that a sample size of at least 9 participants is required. For this study, 10 participants were used in order to have an equal number of females and males. This size was in agreement with several existing biomechanics reliability studies which range from 8-12 subjects (Aaras et al., 1996; Shen et al., 1997; Peach, 1998; Rainoldi et al., 1999; Stokdijk, 2000; Rainoldi et al., 2001; Falla et al., 2002).

For this research, participation was limited to right-hand dominant individuals (due to the limitations of the simulated task design), who either have had recent manual work experience or who performed upper extremity exercise on a regular basis. Cicchetti (1999) reports that reliability levels are influenced by the type of participant being assessed. If the participant is not familiar with manual work or does not exercise their upper extremity regularly, reliability levels obtained from this research could misrepresent the target population of automobile plant workers. Potential participants were also be screened for any recent injuries or musculoskeletal disorders that might affect their performance in the experiment (Appendix B: Form B). Participants were compensated at $10/hour.

3.3 Experimental Task and Equipment

A single simulated overhead task was designed and implemented in the laboratory. The tool mass was 1.25 kg and working cycle time was 54 seconds, consistent with an automobile plant's assembly tasks. A duty cycle of 50% was used here, involving 27 seconds of both work and rest.

The simulated overhead task required participants to stand underneath a height-adjustable (up to 263cm) overhead platform (Figure 4). A keyboard was attached to the bottom of this platform. Participants used a common (non-functional) electric drill, into which a short wooden dowel was chucked, to tap four designated buttons on the keyboard in sequence. Three thin wires were strung over the keyboard, which required the participants to move the drill vertically to move between keys. The task, including precise movements to targets and obstacle avoidance, was designed based on observations of several overhead tasks in an automobile facility and was considered representative of typical task demands.

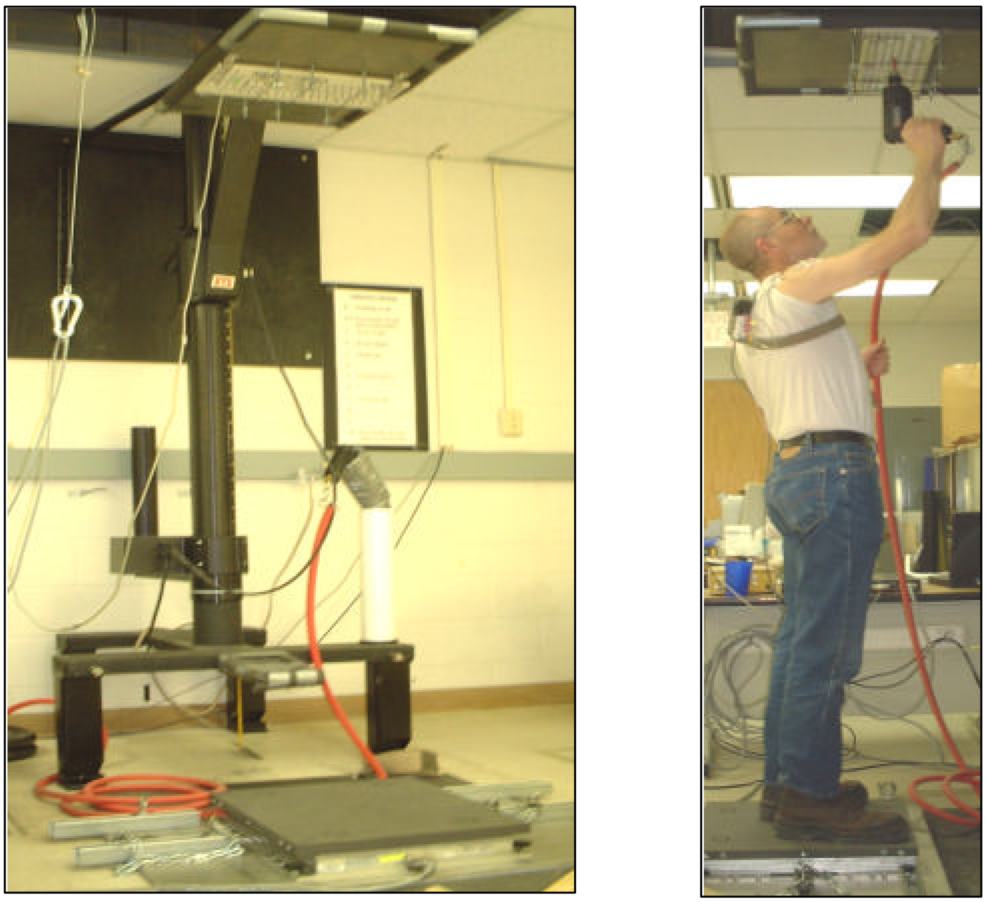


Figure 4: Workstation setup, and participant performing task at middle height.

Participants tapped keys at a fixed paced, 80 beats/min, which was set by a digital metronome. This pace was selected, because it was slow enough that participants could adapt to the speed, but fast enough so that the task did not have a substantial static component. Participants were instructed that keeping pace with the metronome is more important than the accuracy of hitting the correct keys, but that they should also do their best to hit the correct keys in the correct order.

Prior to starting each work cycle, a warning beep was given to allow the participant to prepare for the task (take tool in right hand and hold cable in left hand). The participant started the work cycle after hearing a second start beep. A stop beep indicated the end of duty cycle, at which time the participant placed the drill in a small holder. During rest periods, the participant took several steps to the right side of the workstation and performed light manual work involving screwing or unscrewing nuts and bolts. A warning beep was then provided to indicate there will be 5 seconds before the next work cycle was to begin. Each of these elements was carried out until 60 minutes had elapsed.

3.4 Independent and Dependent Variables

For this study, the independent variable was time. Test-retest reliability involves taking measurements using the same experimenter, from the same group of subjects, with the same instruments, under equivalent conditions, but at different times. This study explored between-days reliability.

The dependent variables were measures of MVE, EMG, and subjective ratings of discomfort. As fatigue indicators, RMS, MnPF, and MdPF intercept and slope were collected for each muscle. EMG was collected using an EMG amplifier (Measurement Systems, Inc.) with a bandpass filter of 10-500Hz, a preamplifier with a gain of 100, and sampled at 2048Hz with Labview 5.1. A Butterworth filter with a cutoff frequency of 3 Hz was applied to the data. Increases in RMS and decreases in MnPF/MdPF are typically concurrent with muscle fatigue (Rainoldi, 1999). Reported EMG parameters included intercepts and slopes for MnPF, MdPF, and RMS determined from linear regression with respect to time. The Borg CR-10 scale (Borg,

1982) was used to evaluate subjective pain and discomfort. Reported ratings of perceived discomfort (RPD) parameters included RPD slope and RPD final rating.

3.5 Experimental Procedures

3.5.1 IRB, Anthropometrics, and Demographics

Upon arrival, participants were introduced to the experiment and signed a written informed consent document, using procedures approved by the Virginia Tech Institutional Review Board (IRB), and included in Appendix A. Anthropometric measurements were taken: weight, stature, shoulder height (measured from floor to acromion), upper arm length (measured from the acromion to lateral epicondyle with arm held horizontally and in the frontal plane), lower arm length (measured from the lateral epicondyle to the wrist crease with the arm held horizontally and in the frontal plane), hand height at full arm extension overhead (with shoulders parallel to ground), and hand height with the elbow at 90^o and the upper arm horizontal (Appendix B). Hand height measurements were taken at the center of the hand grip. Except for weight and stature, all anthropometric data were collected from the right arm. All the measurements were taken with participants wearing shoes, as work on automobile assembly lines requires safety shoes/boots. To guard against any intra-subject variability arising from different shoe heights, participants were asked to wear the same shoes for all experimental sessions. Additionally, demographic information was collected regarding each participant's level of fitness, manual work experience (type and duration of manual work), and any previous injuries (Appendix B).

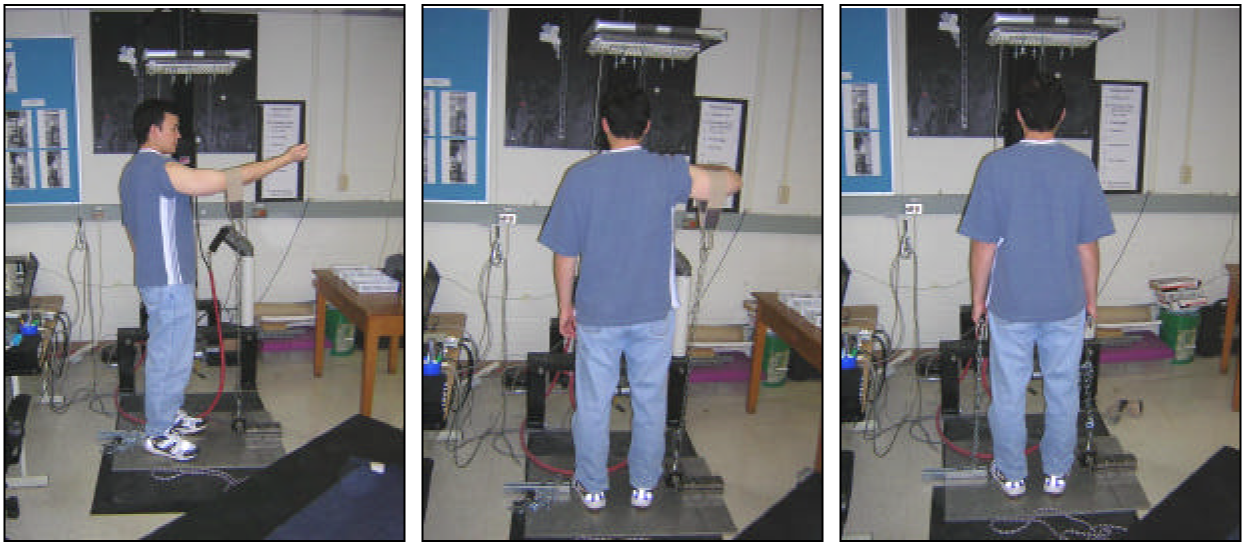
3.5.2 Electrode Placement and Preparation

Electrode pairs were placed on the participant's three shoulder muscles as follows: anterior deltoid (AD) – midway between the lateral third of the clavicle and the deltoid insertion, middle deltoid (MD) – midway between the acromion and the deltoid insertion, and descending trapezius (TRAP) – 2cm lateral to the midpoint of a line connecting the C7 spinous process and the acromion (Nussbaum, 2001). These muscles were selected as having evidence of fatigue during overhead work (Nussbaum, 2001; Kim et al., 2003) and are accessible using surface electrodes. The reference (ground) electrode was placed on the clavicle. Resistance within each electrode pair was kept below 10k? , and electrode distances from the acromion were measured for consistency between sessions (Appendix C).

3.5.3 Maximum Voluntary Exertions

Maximum voluntary exertions (MVE) using the right hand was performed to allow for estimation of strength and maximum electromyographic (EMG) activity levels for the three shoulder muscles. Peak EMG values were obtained during these MVEs in order to normalize EMG data obtained during the experimental trials.

The static postures used to isolate specific shoulder muscles are shown in Figure 5. For the anterior and middle deltoid, participants pulled against a harness with their right arm. For the trapezius, participants pulled against two hand grips. Forces were obtained by a load cell for the right arm. Force and EMG data was collected for a span of 6 seconds. MVE force data was collected with a three-axis load cell (AMTI model MSA-6-250) and sampled a 2048Hz with Labview 5.1. During this time, participants were asked to ramp up to their maximum, hold and ramp down. These MVEs were similar to those performed in a study by Nussbaum et al. (2001).



Anterior Deltoid Middle Deltoid Trapezius

Figure 5: Static MVE postures for the anterior deltoid, middle deltoid, and trapezius.

Participants performed MVEs for at least three trials per muscle with a minimum of 1 minute rest between trials. If the third MVE trial exhibited the highest force for a particular muscle, then the participant continued to perform MVEs. If the subsequent trials were greater in force, then the participant continued to perform more MVEs until a following trial is less than the max.

3.5.4 Understanding RPD

Participants were asked to hold a mass of .5kg with their non-dominant arm parallel to the ground, with their palm facing the floor, in the coronal plane until they reached their endurance limit. During this time, they were asked to recite the Borg scale out loud proceeding from 1 to 10. After this exercise, it was explained that the level of 10 will be very close to the point when they stop the exercise due to extreme discomfort. If participants reached 10 too early, or too late, the Borg scale then was explained again and they were told to adjust their interpretation of the Borg scale.

3.5.5 Performing the Task

The task height of the experiment was adjusted to the participant's anthropometry as described above. After performing MVEs, each participant practiced at the determined height for four minutes. This was needed to warm-up the muscles and to ensure the participant understood the Borg CR-10 Scale. The participant rested at least 5 minutes or until the participant reached a minimal Rating of Perceived Discomfort (RPD) close to 0. The overhead task was then performed for one hour. Subjective RPDs based on the Borg CR-10 were also obtained at the end of every fifth cycle (starting with the second cycle).

3.6 Analysis

Statistics were calculated using the statistical package JMP (version 5). For test-rest reliability indexes, ICC(2,1), SEM, and CV were reported. ICC, used to measure relative reliability, was interpreted using the following scheme: 0-.39 poor, .40-.59 fair, .60-.79 good, .8-1.0 excellent (as described in Chapter 1). SEM, used to measure absolute reliability, was reported along with ICC to account for cases wherein ICC overestimates reliability. CV was reported to help compare different parameters and scales, such as the reliability of EMG parameters to the RPD parameters. Additionally 95% confidence intervals were reported (equation 4) for parameters with high reliability.

Correlations between the RPD slope and all EMG slopes were determined. Spearman's rho was used because of the non-normal distribution of EMG slopes. Correlations were significant at $p < .05$.

CHAPTER 4. RESULTS

4.1 Participants

Five female participants and five male participants completed the procedures described above. Right hand dominance was self-reported by all. Their age ranged from 20 to 27 years of age (mean = 24.20; median = 24.50; sd = 2.70). The mean length of reported employment was 1.8 years (median = .9; sd = 1.8). Only one participant did not have manual work experience, but this participant had performed intense physical activity for 10 years and appeared extremely fit.

All of the 10 participants had average or above average levels of general fitness, based on their reported fitness and typical daily levels of physical exertion. The nine participants who had done manual work had all performed some type of heavy lifting task. None of the participants reported any musculoskeletal problems that might have impeded their performance on the experimental task. Table 2 provides a summary of the participants' anthropometric data.

Table 2: Age and anthropometric data from 10 participants.

				Percentiles	
Measure	**Mean**	**Median**	**SD**	**5th**	**95th**
Age (years)	24.20	24.50	2.70	20.00	27.00
Weight (kg)	72.52	69.95	10.02	64.94	88.84
Stature (cm)	175.74	176.00	5.71	168.68	184.34
Shoulder Height (cm)	145.88	144.80	5.95	138.97	154.75
Upper Arm Length (cm)	30.59	30.40	2.87	26.45	34.64
Lower Arm Length (cm)	25.34	25.75	2.72	20.95	28.79
Arm in full extension (cm)	201.56	202.50	14.94	179.78	218.13
Arm at 90degrees (cm)	175.68	173.50	9.71	166.29	192.46
Working Height (cm)	186.03	184.90	7.17	178.97	197.49

4.2 Maximum Voluntary Exertions (MVE)

Means, ICC, SEM, and CV are presented for MVEs in Table 3. The ICCs for the anterior deltoid, middle deltoid, and trapezius were .96, .95, and .97, respectively. These high values indicate excellent reliability. The low SEM and CV values concur with this interpretation of reliability.

Table 3: MVE means, test-retest ICC, SEM, and CV for each muscle.

Muscle	Mean (N)	ICC	SEM	CV (%)
Anterior Deltoid	165.08	0.96	11.87	7.19
Middle Deltoid	214.86	0.95	16.01	7.45
Trapezius	655.92	0.97	27.12	4.13

4.3 Ratings of Perceived Discomfort (RPD)

Borg scale ratings of perceived discomfort changed in a roughly linear manner with time (Figure 6). All participants reported an RPD of 0 before starting the task. As none of the participants reached their endurance limit during the one hour session, none of the final RPDs reached 10.

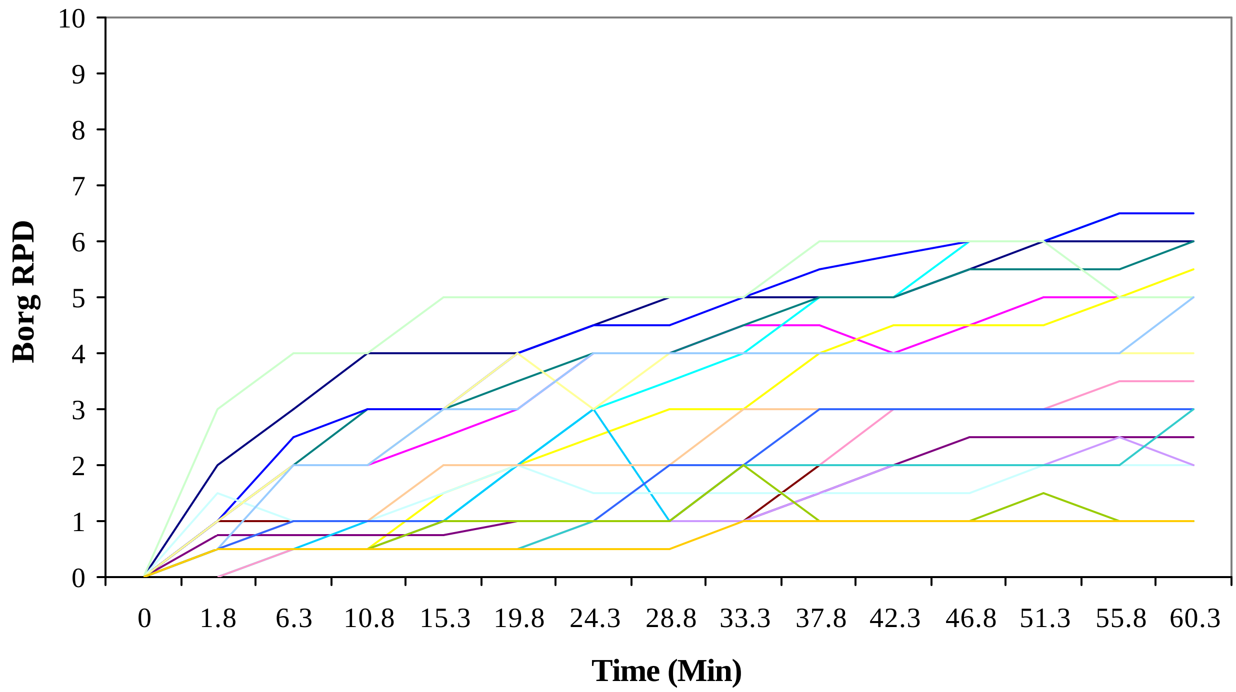


Figure 6: RPD as a function of time.

Means, ICC, SEM, and CV are presented for Borg parameters in Table 4. The ICC for the RPD final value and slope are .86 and .92 respectively. These high values indicate excellent reliability. The low SEM and CV values concur with this interpretation.

Table 4: RPD means, test-retest ICC, SEM, and CV for each muscle for each parameter.

Parameter	Mean	ICC	SEM	CV (%)
Final Rating	3.92	0.86	.77	19.55
Slope (Δ/min)	.047	.92	.0079	16.93

4.4 Electromyography (EMG)

4.4.1 Mean/Median Power Frequency (Mn/MdPF)

MnPF and MdPF trends for the anterior deltoid, middle deltoid, and trapezius were roughly linear with respect to time. Figure 7 shows the Mn/MdPF of one participant for one hour.

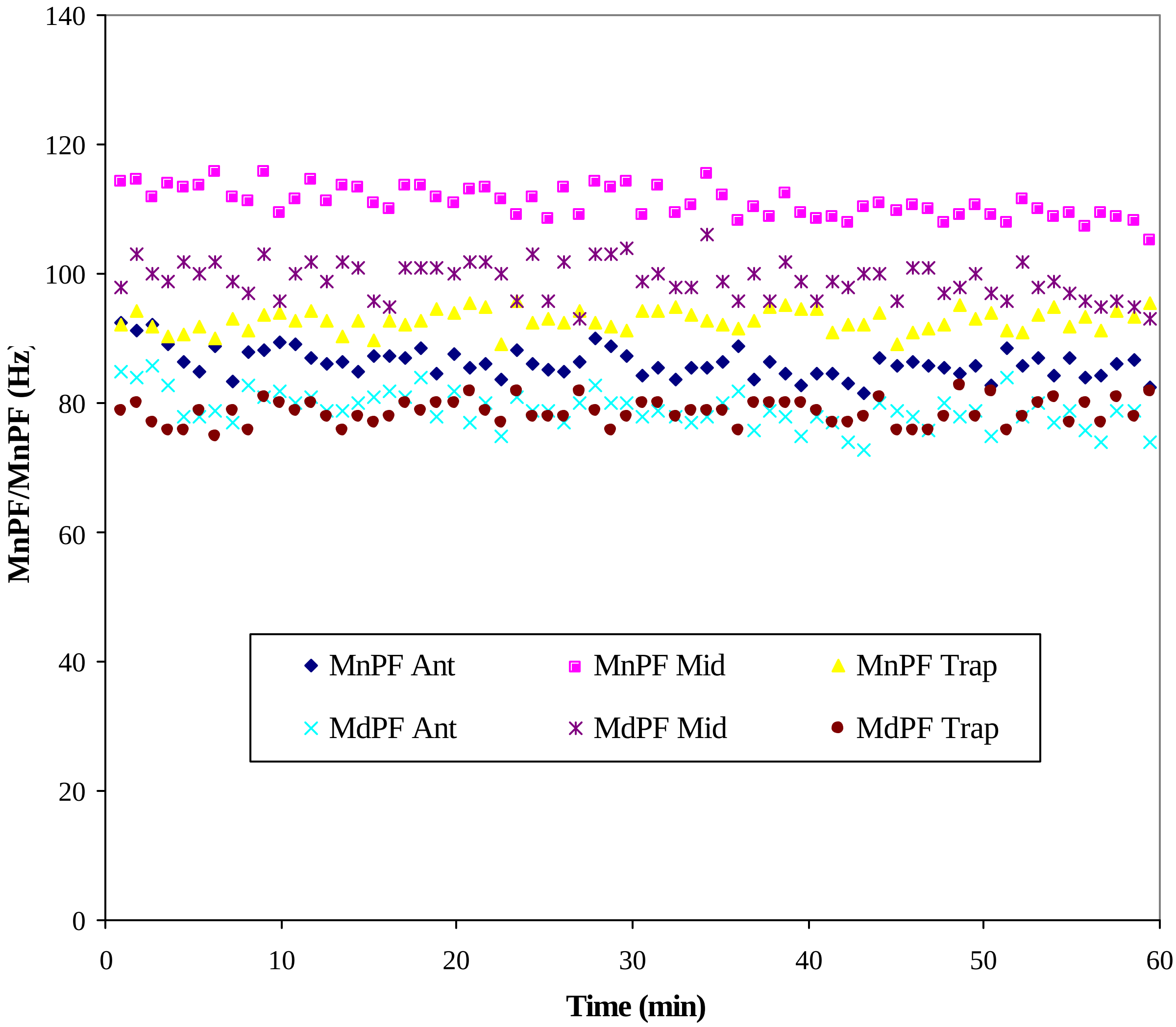


Figure 7: MnPF and MdPF as a function of work cycle for one participant.

Means, ICC, SEM, and CV were presented for Mn/MdPF intercepts in Table 5. The ICC for the MnPF and MdPF intercepts of the anterior deltoid are .79 (good reliability) and .63 (good reliability), respectively. The ICC for the MnPF and MdPF intercepts of the middle deltoid were .81 (excellent) and .83 (excellent), respectively. The ICC for the MnPF and MdPF intercepts of the trapezius were .48 (fair) and .72 (good), respectively. Except for the MdPF intercept of the anterior deltoid, the low SEM and CV showed agreement with the high ICC values.

Table 5: Mn/MdPF intercept means, test-retest ICC, SEM, and CV for each muscle.

Parameter	Muscle	Mean (Hz)	ICC	SEM	CV (%)
MnPF Intercept	Anterior	92.02	0.79	7.17	7.79
	Middle	98.53	0.81	5.60	5.69
	Trapezius	100.68	0.48*	7.88	7.82
MdPF Intercept	Anterior	81.79	0.63	8.54	10.45
	Middle	86.03	0.83	4.99	5.80
	Trapezius	92.73	0.72	5.79	5.75

* Disagreement between ICC and SEM/CV

Means, ICC, SEM, and CV were presented for Mn/MdPF actual slopes and normalized actual slopes in Table 6. (Note: The term "actual" slope is used in order to distinguish this parameter from the "absolute" slope parameter wherein negative signs are removed. Some researchers use the term absolute to indicate non-normalized slope, which is not the case here). The ICC for the MnPF actual slope for the three muscles was 0.0 (poor), .72 (good), and .64 (good), respectively. The ICC for the MnPF normalized actual slope for the three muscles are 0 (poor), .72 (good), and .59 (fair), respectively. Both parameters, however, showed extremely high SEM and CV. Both parameters only showed agreement between ICC and SEM/CV for the anterior muscle.

Reliability statistics for MdPF actual slope and normalized actual slope (Table 6) showed similar trends to MnPF. The ICC for the actual slope for the three muscles was 0.0 (poor), .80 (excellent), and .68 (good), respectively. The ICC for the normalized actual slope for the three muscles was .12 (poor), .79 (good), and .64 (good), respectively. Both parameters showed extremely high SEM and CV. Like MnPF, only the anterior muscle for both parameters showed agreement between ICC and SEM/CV.

Table 6: Mn/MdPF actual slope and normalized actual slope means, test-retest ICC, SEM, and CV for each muscle.

Parameter	Muscle	Mean (Hz/cycle)	ICC	SEM	CV (%)
Actual MnPF Slope	Anterior	-.035	0@	0.070	203.43
	Middle	.016	0.72*	0.063	385.23
	Trapezius	-.017	0.64*	0.032	183.04
Normalized Actual MnPF Slope	Anterior	-.00037	0@	0.00077	207.60
	Middle	.00018	0.72*	0.00062	352.29
	Trapezius	-.00016	0.59*	0.00033	202.25
Actual MdPF Slope	Anterior	-.036	0@	0.057	160.33
	Middle	.024	0.80*	0.054	222.87
	Trapezius	-.016	0.68*	0.029	205.90
Normalized Actual MdPF Slope	Anterior	-.00044	0.12	0.00070	156.85
	Middle	.00027	0.79*	0.00061	225.16
	Trapezius	-.00015	0.64*	0.00036	229.94

* Disagreement between ICC and SEM/CV
@ Negative or 0 ICC value

Means, ICC, SEM, and CV are presented for Mn/MdPF absolute slopes and normalized absolute slopes in Table 7. The ICC of the absolute slope for the three muscles was .3 (poor), .26 (poor), and .18 (poor), respectively. The ICC of the normalized absolute slope for the three muscles was .94 (excellent), .49 (fair), and .13 (poor), respectively. The SEM and CV levels concurred with this interpretation.

The ICC for the MdPF absolute slope for the three muscles was .69 (good), .57 (fair), and .28 (poor), respectively. The ICC for the MdPF normalized absolute slope for the three muscles was .20 (poor), .26 (poor), and .032 (poor), respectively. The SEM and CV levels concurred with this interpretation.

Table 7: Mn/MdPF absolute slope and normalized absolute slope means, test-retest ICC, SEM, and CV for each muscle.

Parameter	Muscle	Mean (Hz/cycle)	ICC	SEM	CV (%)
Absolute MnPF Slope	Anterior	.061	0.30	0.035	57.40
	Middle	.095	0.26	0.073	80.86
	Trapezius	.045	0.18	0.028	61.30
Normalized Absolute MnPF Slope	Anterior	.00067	0.94	0.00012	17.19
	Middle	.00095	0.49	0.00047	49.83
	Trapezius	.00045	0.13	0.00027	62.26
Absolute MdPF Slope	Anterior	.054	0.69	0.022	40.38
	Middle	.090	0.57	0.056	61.75
	Trapezius	.048	0.28	0.028	59.21
Normalized Absolute MdPF Slope	Anterior	.00069	0.20	0.00046	66.79
	Middle	.00010	0.26	0.00075	74.13
	Trapezius	.00050	0.032	0.00085	84.79

4.4.2 Root Means Square (RMS)

Means, ICC, SEM, and CV are presented for RMS intercepts, actual slopes and absolute slopes in Table 8. The ICC for the RMS intercept for the three muscles was .79 (good), .38 (poor), and .60 (good), respectively. The ICC for the RMS actual slope for the three muscles was .64 (good), .78 (good), and .27 (poor), respectively. The ICC for the RMS absolute slope was .34 (poor), .91 (excellent), and .030 (poor). The SEM and CV levels showed disagreement with corresponding ICC values for all parameters but the RMS intercept for the anterior deltoid and trapezius, and the RMS absolute slope for the anterior deltoid.

Table 8: RMS intercept, actual slope, and absolute slope means, test-retest ICC, SEM, and CV for each muscle.

Parameter	Muscle	Mean	ICC	SEM	CV (%)
RMS Intercept	Anterior	0.21	0.79	0.037	17.48
	Middle	0.17	0.38 [x]	0.057	33.53
	Trapezius	0.15	0.60	0.039	25.88
Actual RMS Slope	Anterior	6.23E-05	0.64 [x]	0.00026	411.67
	Middle	3.21E-04	0.78 [x]	0.00026	80.87
	Trapezius	-1.20E-05	0.27 [x]	0.00030	2479.73
Absolute RMS Slope	Anterior	3.03E-04	0.34	0.00024	79.85
	Middle	4.14E-04	0.91 [x]	0.00015	35.93
	Trapezius	3.04E-04	0.030 [x]	0.00016	52.00

[x] Disagreement between ICC and SEM/CV

4.4.3 EMG Reliability Comparison

Table 9 provides the ICC reliability classification for the Mn/MdPF parameters of intercept, actual slope, normalized actual slope, absolute slope, normalized absolute slope for each muscle. Table 9 also provides the ICC reliability classification for the RMS parameters of intercept, actual slope, and absolute slope for each muscle. Table 10 provides the CV for Mn/MdPF and RMS parameters.

Amongst the three muscles, the middle deltoid appeared to have more reliable parameters. Both MnPF and MdPF intercept parameters had good to excellent reliability. The RMS intercept parameter showed good reliability for both anterior and trapezius muscles. Normalizing the actual slope did not seem to increase reliability, and in some cases decreased reliability. Except for the anterior deltoid, MnPF and MdPF absolute slope parameters showed less reliability than either actual slope or normalized actual slope. Mn/MdPF normalized absolute slope showed similar reliability trends as Mn/MdPF absolute slope. RMS absolute slope did not show dramatic changes in reliability from actual slope.

Table 9: ICC Reliability Classification for MnPF, MdPF, and RMS parameters.

EMG	Parameter	Anterior Deltoid	Middle Deltoid	Trapezius
MnPF	Intercept	Good	Excellent	Good [x]
	Actual Slope	Poor	Good [x]	Good [x]
	Normalized Actual Slope	Poor	Good [x]	Fair [x]
	Absolute Slope	Poor	Poor	Poor
	Normalized Absolute Slope	Excellent	Fair	Poor
MdPF	Intercept	Good	Excellent	Good
	Actual Slope	Poor	Excellent [x]	Good [x]
	Normalized Actual Slope	Poor	Good [x]	Good [x]
	Absolute Slope	Good	Fair	Poor
	Normalized Absolute Slope	Poor	Poor	Poor
RMS	Intercept	Good	Poor [x]	Good
	Actual Slope	Good [x]	Good [x]	Poor [x]
	Absolute Slope	Poor	Excellent [x]	Poor [x]

[x] Disagreement between ICC and SEM/CV

Amongst the three muscles, no single one was evident as having the most reliable (low CV percentage) parameters. Both MnPF and MdPF intercept parameters showed low CVs indicating high reliability. Though slightly higher than MnPF/MdPF, the RMS intercept parameter showed relatively low CV. Normalizing the actual slope did seem to make a noticeable increase in reliability. Taking the absolute slope decreased the CV of the Mn/MdPF actual slope by at least a factor of 3. Mn/MdPF normalized absolute slope showed similar trends to absolute slope. RMS absolute slope did not show dramatic changes in reliability from actual slope. Taking the absolute slope decreased the CV of the RMS actual slope by at least a factor of 2.

Table 10: CV Percentage for MnPF, MdPF, and RMS parameters.

EMG	Parameter	Anterior Delt.	Middle Delt.	Trapezius
MnPF	Intercept	7.79	5.69	7.82
	Actual Slope	203.43	385.23	183.04
	Normalized Actual Slope	207.6	352.29	202.25
	Absolute Slope	57.4	80.86	61.3
	Normalized Absolute Slope	17.19	49.83	62.26
MdPF	Intercept	10.45	5.8	5.75
	Actual Slope	160.33	222.87	205.9
	Normalized Actual Slope	156.85	225.16	229.94
	Absolute Slope	40.38	61.75	59.21
	Normalized Absolute Slope	66.79	74.13	84.79
RMS	Intercept	17.48	33.53	25.88
	Actual Slope	411.67	80.87	2479.73
	Absolute Slope	79.85	35.93	52

4.5 EMG and RPD correlation

Due to the non-normal distribution of the EMG slopes, Spearman's rho is reported for the correlation between each EMG parameter slope and the Borg slope (Table 11). The middle deltoid showed slightly higher correlation levels for Mn/MdPF for actual and normalized slopes. RMS slopes showed higher correlation levels than MnPF and MdPF slopes. There were no significant correlations for RMS.

Table 11: Spearman's Rho for Each Muscle at Each EMG Parameter correlated with Borg slope

EMG	Parameter	Muscle	Spearman's rho	p-value
MnPF	Actual Slope	Anterior Deltoid	0.0782	0.743
		Middle Deltoid	0.298	0.202
		Trapezius	-0.009	0.970
	Normalized Actual Slope	Anterior Deltoid	0.0376	0.875
		Middle Deltoid	0.269	0.251
		Trapezius	-0.0135	0.955
	Absolute Slope	Anterior Deltoid	0.0932	0.696
		Middle Deltoid	-0.0316	0.895
		Trapezius	0.0571	0.811
	Normalized Absolute Slope	Anterior Deltoid	0.153	0.519
		Middle Deltoid	0.003	0.990
		Trapezius	-0.015	0.950
MdPF	Actual Slope	Anterior Deltoid	0.0027	0.729
		Middle Deltoid	0.263	0.262
		Trapezius	-0.0026	0.925
	Normalized Actual Slope	Anterior Deltoid	0.0662	0.782
		Middle Deltoid	0.280	0.232
		Trapezius	-0.0165	0.945
	Absolute Slope	Anterior Deltoid	0.239	0.310
		Middle Deltoid	-0.301	0.900
		Trapezius	0.0647	0.787
	Normalized Absolute Slope	Anterior Deltoid	0.277	0.248
		Middle Deltoid	0.0045	0.985
		Trapezius	0.0241	0.920
RMS	Actual Slope	Anterior Deltoid	-0.412	0.071
		Middle Deltoid	0.274	0.243
		Trapezius	-0.387	0.092
	Absolute Slope	Anterior Deltoid	-0.114	0.631
		Middle Deltoid	0.289	0.217
		Trapezius	-0.203	0.391

* Significant, $p < .05$

CHAPTER 5. DISCUSSION

The goal of this research was to determine the reliability of fatigue measures in a dynamic intermittent overhead work task. To estimate reliability, indexes of ICC, SEM, and CV were used. Fatigue measures included RPD (final rating and slope parameters), and EMG (intercept and various slope parameters). Parameters with high reliability indicated precise results and consistent methods. High ICC and low SEM/CV were indicators of this.

5.1 Maximum Voluntary Exertions (MVE)

MVEs for all three muscles showed excellent reliability (ICC of .95-.97, CV of 4.13 to 7.45 %). Others studies found similar levels of reliability for MVE. Elfving et al. (1999) reported an ICC of .93 and CV of 10.7% for an MVE involving maximum trunk extension. Rainoldi et al. (2001) reported CVs of 1.1% to 6.4% for an MVE involving leg extensions. Overall, the MVE methods for this study would thus appear to be reliable.

5.2 Ratings of Perceived Discomfort (RPD)

RPD parameters indicated relatively high reliability. The RPD final rating showed similar reliability characteristics (ICC of .86, SEM of .77, and CV of 19.55%) to another study. Elfving et al. (1999) reported an ICC of .84, SEM of .8, and CV of 17% for a task requiring a 45-sec isometric contraction.

Figure 6 indicated the linear nature of the RPD ratings over time. This study suggests that both slope and final rating of Borg CR-10 scale for discomfort are reliable parameters. Nevertheless, caution should be taken when using final reported values. CV indicates that the true final value may be within ˜20% of the reported final value. Thus, if a participant reports a final value of 1 then the actual value lies within .8 to 1.2, however, if the reported final value is

10 then the actual value lies within 8 to 12 (in Borg 10-point scale, range is limited from 8 to 10). Thus, with higher final reported values, more absolute uncertainty of the true final value exists.

By utilizing SEM (Equation 4), 95% confidence intervals of slope and final rating were determined. These are ±.018 and ±1.74, respectively. Variation within these ranges should be considered normal. If applied to other research or the industrial setting, any variation outside of these ranges (above upper limit) may indicate the possibility for injury.

It is recommended that 95% confidence intervals based on SEM only be used when the parameters are determined to be highly reliable. In the case below, wherein EMG slopes show lower levels of reliability, a confidence interval would not be useful here. The result would be a very wide confidence interval that, if applied to EMG fatigue research, may mislead researchers as to what values of slopes are reliable.

5.3 Electromyography (EMG)

The results for MnPF, MdPF, and RMS parameters indicate mixed levels of reliability depending on whether reliability is interpreted from ICC or CV. This disagreement creates some difficulty in interpreting the reliability of EMG parameters. Out of the number of reported parameter ICCs for EMG, 5 of 15 for MnPF, 4 of 15 for MdPF, and 6 of 9 for RMS showed reliability disagreement with CV. From another perspective, out of the number of reported ICCs for the three muscles, 1 of 13 for anterior deltoid, 7 of 13 for middle deltoid, and 7 of 13 for trapezius showed reliability disagreement with CV. If disagreement between ICC and CV is an indication of poor reliability, it could be said that for this study RMS parameters have worse reliability than MnPF or MdPF parameters. Likewise, the parameters of the anterior deltoid muscle have the best reliability of the three muscles.

This situation of conflicting ICC and CV is important in reliability studies, however there are few studies which show any preference between the two. This issue is further discussed in section 5.5 (see below). In the following sections, ICC and CV results for EMG are first interpreted separately, and then reliability conclusions about the parameters are made after comparing ICC and CV. Additionally, any reliability conclusions inferred from the following sections are specific to the muscles and task studied.

5.3.1 MnPF and MdPF Intercept

ICCs for MnPF intercepts showed fair to excellent reliability (.48-.79) and MdPF intercept ICCs showed good to excellent levels of reliability (ICC of .63-.83). This is in agreement with other studies. Larsson et al. (2003) reported ICCs of .53-.88 for MnPF intercept of different leg muscles. Elfving et al. (1999) reported ICCs of .41 to .70 for MdPF intercept, while Mannion et.al (1994) reported ICCs of .70 to .77 of different back muscles.

CVs were fairly low for all MdPF and MdPF intercepts (5.69 to 7.82%, and 5.75 to 10.45% respectively). Other studies also showed low CVs. A study by Larrson et al. (2003) showed CVs of 5.9% to 10.2% (estimated from authors' SEM and means), and Elfving et al. (1999) reported CVs of 8.2% to 10.2%. These low CVs indicate corresponding ICC should result in good to excellent reliability.

The ICCs for MnPF and MdPF intercepts of the anterior and middle deltoid are similar and correspond to the CVs. Therefore, the low ICC of .48 for the MnPF intercept of the trapezius may be an outlier, possibly a case wherein ICC underestimates reliability. Poor precision in electrode placement and replacement (over test and retest days) can result in lower levels of reliability for MnPF or MdPF intercepts (Larsson et al., 2003). However, since the MdPF intercept for trapezius was low (7.82%) and the MdPF intercept showed an ICC of .72, indicating a trend for higher ICCs, electrode placement is probably not an issue here. Overall, MnPF and MdPF intercepts are both highly reliable parameters.

The 95% confidence intervals of MnPF intercept for anterior deltoid, middle deltoid, and trapezius are ±16.2Hz, ±12.7Hz, and ±17.8Hz, respectively. The 95% confidence intervals of MdPF intercept for anterior deltoid, middle deltoid, and trapezius are ±19.3 Hz, ±11.3Hz, and ±13.1Hz, respectively. Variation within these ranges should be considered normal. Intercepts reported outside these ranges may be indication of inconsistent EMG methods, such as electrode placement.

5.3.2 MnPF and MdPF Slopes

Actual slopes and normalized actual slopes for MnPF and MdPF indicated a wide range of reliability levels. The ICC of the anterior deltoid for these parameters showed poor reliability (<0 to .12). The ICC of the middle deltoid and trapezius showed fair to excellent levels of reliability (.59 to .80). These mixed levels of reliability agree with another study, wherein Elfving et al. (1999) reported MdPF ICCs of .04 to .45. In this study, normalizing the slope did not seem to improve reliability. Likewise, Dedering et al. (2000) reported MdPF actual slope ICCs of .70 to .87 and not much improved MdPF normalized slope ICCs of .65 to .90 for a task involving back extensor muscles.

CVs ranged from 156.85% to 385.23%. These values are much higher than other studies. Elfving et al. (1999) reported MdPF actual slope CVs of 35% to 75%, while Dedering et al. (2000) reported CVs of 14.4% to 26.2%. Dedring et al. (2000) also reported normalized slope CVs of 12% to 23.9%, which are slightly more reliable than actual slopes. In this study, there was no obvious trend of whether normalization improved slope reliability for specific muscles or whether MnPF was more or less reliable than MdPF. Lariviere et al. (2002) suggests that normalization generally decreases inter-subject variability and may decrease reliability. For this task, actual slopes and normalized actual slopes for MnPF and MdPF are parameters with poor reliability.

The ICCs for the middle deltoid and trapezius (Mn/MdPF actual and normalized slope) were suspect, because of their disagreement with corresponding CV values and with the ICC for anterior deltoid. The occurrence of both negative and positive slopes is most likely the cause for disagreement. For actual slope, the anterior deltoid (Table 12) had the most negatives and also had the most within-subject mix of negatives and positives. These two factors may have caused an inflation of the within and between means square error resulting in a near 0 or negative ICC. On the other hand, both the middle deltoid and trapezius had fewer negative slopes and fewer within-subject mixes of negatives and positives. These two factors may have caused an inflated between means square error, resulting in a relatively high, but overestimated ICC. Additionally, MdPF middle deltoid and trapezius ICCs are slightly higher than MnPF middle deltoid and trapezius ICCs, because of less within-subject mix of signs, resulting in a relatively higher between-subject error.

Table 12: MnPF and MdPF Actual Slopes on Day 1 (Test) and Day 2 (Re-test) Across Muscles.

Days	MnPF			MdPF		
	Anterior	Middle	Trapezius	Anterior	Middle	Trapezius
1	-0.087	-0.011	-0.049	-0.102	0.017	-0.023
2	0.176	-0.061	-0.069	0.113	0.010	-0.052
1	-0.072	0.122	0.062	-0.046	0.193	0.090
2	0.024	0.322	0.057	0.034	0.372	0.064
1	-0.044	-0.071	-0.097	-0.031	-0.062	-0.105
2	-0.040	-0.024	-0.083	-0.026	-0.004	-0.083
1	-0.051	-0.075	-0.049	-0.048	-0.076	-0.046
2	-0.111	-0.015	-0.016	-0.079	-0.006	-0.020
1	-0.044	0.050	0.102	-0.008	0.058	0.104
2	-0.062	0.088	0.016	-0.032	0.113	0.036
1	-0.018	0.027	-0.0067	-0.036	-0.0009	-0.0076
2	-0.059	-0.087	-0.021	-0.103	-0.090	-0.023

1	-0.063	0.065	0.024	-0.051	0.037	0.017
2	0.036	0.177	-0.064	0.026	0.106	-0.094
1	0.027	0.152	-0.044	0.011	0.142	-0.044
2	-0.011	0.110	-0.055	-0.003	0.097	-0.060
1	-0.081	-0.121	0.001	-0.072	-0.103	-0.016
2	-0.003	-0.072	-0.061	-0.021	-0.083	-0.064
1	-0.121	-0.156	0.014	-0.130	-0.148	0.0065
2	-0.090	-0.092	-0.007	-0.114	-0.083	0.0060

Shaded boxes indicate within-subject positive and negative slope occurrences

Absolute slope MnPF and MdPF ICCs (.18 to .3 and .28 to .69, respectively) were lower than actual slope MnPF and MdPF ICCs. Absolute slope ICCs tend to agree with corresponding CVs, however and might therefore be more representative reliability values. Normalizing absolute slope for MnPF tended to increase ICCs (.13-.94), while normalizing absolute slope for MdPF tended to decrease ICCs (.032-.26). Given the trend for low ICCs, the .94 value is most likely an outlier, which suggests that normalizing actual slopes for either MnPF or MdPF does not change reliability dramatically.

Absolute slope MnPF CVs (57.4% to 80.86%) were higher than MdPF CVs (40.38% to 61.75%). On the other hand, normalized absolute slope MnPF CVs (17.9% to 62.26%) were lower than MdPF CVs (66.79% to 84.79%). Normalizing MnPF absolute slope tended to increase reliability, while normalizing MdPF absolute slope tended to decrease reliability. Although the reliability differences between MnPF and MdPF are not large, the differences are important. Most authors tend to only report MdPF when considering EMG reliability. When considering CVs, this study indicates that normalized absolute MnPF is more reliable than Mn/MdPF actual slope, Mn/MdPF actual normalized slope, Mn/MdPF absolute slope, and MdPF absolute normalized slope.

The lower CVs and the higher reliability of absolute parameters compared to actual parameters is due to the removal of the negative sign and the reduction of the standard deviation and the mean. Based on the SEM equation, CV will decrease if the standard deviation of the data decreases more in proportion than the mean does. Table 13 provides the ratio of the mean divided by the standard deviation for each parameter. It is evident that this ratio is less for absolute slope and absolute normalized slope than for actual slope and actual normalized slope. Although results may vary, depending on distribution of non-negative versus negative slopes, taking the absolute value of the slope tends to be useful in increasing reliability.

Table 13: Mean/Standard Deviation for Each Parameter Across Each Muscle.

EMG	Muscle	Actual Slope	Actual Normal Slope	Absolute Slope	Absolute Normal Slope
MnPF	Anterior	1.921	2.041	0.687	0.716
	Middle	7.288	6.612	0.748	0.697
	Trapezius	3.033	3.167	0.678	0.667
MdPF	Anterior	1.603	1.674	0.726	0.747
	Middle	5.029	4.896	0.940	0.862
	Trapezius	3.666	3.842	0.697	0.668

5.3.3 RMS Intercepts and Slopes

RMS intercept ICCs (.38 to .79) indicated poor to good reliability. The poor reliability of the middle deltoid (.38) is most likely an underestimation of reliability since it does not agree with its corresponding CV or the good reliability of the other two muscles. CVs for all three RMS intercepts indicated relatively good reliability.

The 95% confidence intervals of RMS intercept for anterior deltoid, middle deltoid, and trapezius are ±.084, ±.13, and ±.088, respectively. Variation within these ranges should be considered normal.

RMS actual slope ICCs (.27 to .78) indicated poor to good reliability; however, the higher ICCs may be an over prediction of reliability like ICCs for Mn/MdPF slopes. The mix of negative and positive RMS slope values and the disagreement with corresponding CV values support this. RMS absolute slope (.30 to .91) indicated poor to excellent reliability. Though the wide range of reliability is suspicious, it is difficult to determine how high or low the reliability really is. Thus, CVs will be used to interpret RMS slopes.

CVs for RMS slope were very high (80.87% to 2479.73%) indicating very low reliability. Taking the absolute value of RMS slope did increase reliability (35.93% to 79.85%). RMS absolute slope is comparable to MnPF absolute normalized slope in terms of CVs. Of all the slope parameters, MnPF absolute normalized slope and RMS absolute slope were the most reliable.

In comparing EMG slopes, results for this study showed some compliance with a study by Larviere et al. (2002), which reported that MdPF slope was more reliable than RMS slope. Based on CVs of the current study, this is true for the anterior deltoid and the trapezius. Larviere et al. (2002) suggests that RMS slope is more influenced by load than MnPF slope. The slight disagreement with this study may be due to this task having a light load overall (weight of arm plus tool mass), where in the cited study, participants performed at higher loads with static trunk extensions at 75% MVC.

The reason for low reliability of slopes is not clear. Elfving et al. (1999) suggests that some uncontrollable factors such as metabolite production and vascular flow may affect the reliability of slopes. The dynamic nature of this task requiring the use of multiple muscles may also cause low reliability. Attributes of a dynamic task that might affect EMG reliability include: changes in postures (Nargol et al., 1999), force and moment changes with varying positions and range of motion, movement of the neuromuscular junction with relation to electrode positioning, and problems with non-stationary signals (Larsson et al., 1999). Mannion et al. (1994) reports that a static endurance task, measuring the MdPF slopes for the thoracic and lumbar regions, demonstrated higher reliability when observed for the greater slope of the two muscles rather than either of the individual muscles' slope.

Mannion et al. (1994) commented that the limiting factor for endurance time and changes in EMG are the weaker muscles. For a dynamic task, however, a dominant muscle may try to overcompensate for weaker muscles, which may affect the EMG slope of the less dominant muscles. Additionally, in a static task, it is more difficult to consciously distribute load among different muscles when fatigue occurs, while in a dynamic task load can be easily shifted to supporting muscle groups. This was observed in participants from this study, when some started the task with their upper arm in the sagital plane and as their anterior deltoid fatigued, shifted their arm to the coronal plane to recruit the middle and trapezius muscles. This shifting might result in EMG slope having less sensitivity for individual muscles. It might therefore be advantageous to utilize other parameters or means of measuring reliability for EMG, which considers accounting for all muscles at once instead of individual muscles.

5.4 EMG and RPD correlation

None of the EMG slope parameters were significantly correlated with RPD slopes. There may be a variety of reasons why the correlations were low. One reason is stated above, the recruitment of different muscles and the shifting of posture during dynamic tasks. Combining the slopes by using various methods, such as reported by Mannion et al. (1994) may increase correlation levels. Higher muscle activity while performing the task may also increase correlation levels, as the low overall muscle activity (15%-20% MVE) of the current task might have contributed to low reliability of EMG.

Another reason for the lack of correlation may be that EMG slope is simply not sensitive to this particular dynamic intermittent task. Tables 14, 15, 16, and 17 present p-values for intercepts and slopes (based on a linear regression with respect to time) of MnPF, MdPF, RMS, and RPD, respectively. The number of non-significant slopes were 12 of 60 for MnPF, 18 of 60 for MdPF, 12 of 60 for RMS, 1 of 20 for RPD. Given the non-significant EMG slopes and their low reliability, the problem area likely lies in EMG slopes and not in the RPD.

Table 14: P-values for MnPF Intercept and Slope on Day 1 (Test) and Day 2 (Re-test).

Participant #	Muscle	Day 1 (p-value)		Day 2 (p-value)	
		Intercept	Slope	Intercept	Slope
	Anterior	<.0001	<.0001	<.0001	<.0001
1	Middle	<.0001	0.4721**	<.0001	0.0009
	Trapezius	<.0001	<.0001	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	0.0939**
2	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.0002
	Anterior	<.0001	0.0004	<.0001	0.0004
3	Middle	<.0001	<.0001	<.0001	0.0301
	Trapezius	<.0001	<.0001	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	<.0001
4	Middle	<.0001	<.0001	<.0001	0.4735**
	Trapezius	<.0001	<.0001	<.0001	0.2617
	Anterior	<.0001	0.0212	<.0001	0.0215
5	Middle	<.0001	0.0085	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.1791**
	Anterior	<.0001	0.5357**	<.0001	<.0001
6	Middle	<.0001	0.9373**	<.0001	<.0001
	Trapezius	<.0001	0.0033	<.0001	0.0705**
	Anterior	<.0001	<.0001	<.0001	0.0084
7	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	0.0094	<.0001	0.0006
	Anterior	<.0001	0.0084	<.0001	0.4411**
8	Middle	<.0001	<.0001	<.0001	0.0402
	Trapezius	<.0001	0.0006	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	0.7996**
9	Middle	<.0001	<.0001	<.0001	0.036
	Trapezius	<.0001	0.9121**	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	<.0001
10	Middle	<.0001	0.7361**	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.6372**

** Not significant, p > .05

Table 15: P-values for MdPF Intercept and Slope on Day 1 (Test) and Day 2 (Re-test).

Participant #	Muscle	Day 1 (p-value)		Day 2 (p-value)	
		Intercept	Slope	Intercept	Slope
	Anterior	<.0001	0.0002	<.0001	0.0004
1	Middle	<.0001	0.3172**	<.0001	0.6289**
	Trapezius	<.0001	0.0373	<.0001	0.0001
	Anterior	<.0001	0.0028	<.0001	0.0425
2	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.0003
	Anterior	<.0001	0.0572**	<.0001	0.0304
3	Middle	<.0001	<.0001	<.0001	0.7213**
	Trapezius	<.0001	<.0001	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	<.0001
4	Middle	<.0001	0.0003	<.0001	0.8068**
	Trapezius	<.0001	0.0013	<.0001	0.2572**
	Anterior	<.0001	0.8083**	<.0001	0.0923**
5	Middle	<.0001	0.0072	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.0113
	Anterior	<.0001	0.5617**	<.0001	<.0001
6	Middle	<.0001	0.0167	<.0001	<.0001
	Trapezius	<.0001	0.072	<.0001	0.0732**
	Anterior	<.0001	<.0001	<.0001	0.0603**
7	Middle	<.0001	0.0005	<.0001	<.0001
	Trapezius	<.0001	0.1205**	<.0001	0.0007
	Anterior	<.0001	0.0603**	<.0001	0.8234**
8	Middle	<.0001	<.0001	<.0001	0.0409
	Trapezius	<.0001	0.0007	<.0001	0.0004
	Anterior	<.0001	<.0001	<.0001	0.1788**
9	Middle	<.0001	<.0001	<.0001	0.0464
	Trapezius	<.0001	0.234**	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	<.0001
10	Middle	<.0001	0.7361**	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.6372**

** Not significant, p > .05

Table 16: P-values for RMS Intercept and Slope on Day 1 (Test) and Day 2 (Re-test).

Participant #	Muscle	Day 1 (p-value)		Day 2 (p-value)	
		Intercept	Slope	Intercept	Slope
	Anterior	<.0001	0.0044	<.0001	0.0335
1	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.024
	Anterior	<.0001	0.0001	<.0001	0.6128**
2	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	0.0324
	Anterior	<.0001	<.0001	<.0001	<.0001
3	Middle	<.0001	<.0001	<.0001	<.0001
	Trapezius	<.0001	<.0001	<.0001	<.0001
	Anterior	<.0001	<.0001	<.0001	0.0026
4	Middle	<.0001	<.0001	<.0001	0.0002
	Trapezius	<.0001	<.0001	<.0001	0.0057
	Anterior	<.0001	0.2365**	<.0001	0.7744**
5	Middle	<.0001	0.7622**	<.0001	<.0001
	Trapezius	<.0001	0.0024	<.0001	0.0011
	Anterior	<.0001	0.1427**	<.0001	0.7105**
6	Middle	<.0001	0.0011	<.0001	0.005
	Trapezius	<.0001	<.0001	<.0001	0.0824**
	Anterior	<.0001	<.0001	<.0001	0.0009
7	Middle	<.0001	0.3857**	<.0001	0.001
	Trapezius	<.0001	<.0001	<.0001	0.0003
	Anterior	<.0001	0.1992**	<.0001	0.4668**
8	Middle	<.0001	0.0116	<.0001	0.0004
	Trapezius	<.0001	0.0009	<.0001	0.0079
	Anterior	<.0001	0.0004	<.0001	0.0189
9	Middle	<.0001	<.0001	<.0001	0.0057
	Trapezius	<.0001	0.0052	<.0001	<.0001
	Anterior	<.0001	0.2125**	<.0001	0.0006
10	Middle	<.0001	0.6949**	<.0001	0.02
	Trapezius	<.0001	0.0002	<.0001	0.0034

** Not significant, p > .05

Table 17: P-values for RPD Intercept and Slope on Day 1 (Test) and Day 2 (Re-test).

Participant #	Day 1 (p-value)		Day 2 (p-value)	
	Intercept	Slope	Intercept	Slope
1	<.0001	<.0001	0.0003	<.0001
2	0.3662**	<.0001	0.2899**	<.0001
3	0.1242**	<.0001	0.0012	<.0001
4	<.0001	<.0001	<.0001	<.0001
5	0.01	<.0001	<.0001	0.5554**
6	<.0001	0.0018	0.0004	0.0028
7	0.0008	<.0001	0.085	<.0001
8	0.0363	<.0001	0.0021	<.0001
9	0.0507**	<.0001	0.6289**	<.0001
10	0.0054	0.0198	0.0007	<.0001

** Not significant, $p > .05$

5.5 ICC Versus CV

Of all EMG parameters reported in this study, 15 out of 36 showed disagreement in reliability between ICC and CV. This implies that almost half of all parameters can possibly be misinterpreted. ICC takes into account the between subject variance, but has been known to over and underestimate reliability (Denegar et al., 1993). For this research, given the trends in data, ICC appears to have misrepresented reliabilities of several parameters. This raises the question of whether the ICC index is sufficient or even necessary for determining reliability.

Researchers suggest that ICC and SEM be reported together (Elfving et al., 1999; Larsson et al., 2003). This seems logical, since ICC is a “relative” reliability index, while SEM is an “absolute” reliability index. The main disadvantage of reporting SEM alone is that it cannot be used to compare reliability between different tools or measures. On the other hand, CV is also a “relative” index, which can be used to compare between tools and measures. To the knowledge of this author, CV has not been reported in research to underestimate values for reliability. It may then be more efficient to report only SEM and CV values in future reliability studies.

There are a few more advantages to using CV instead of ICC. When an ICC of 0 to 1 is reported, reliability is interpreted as poor, fair, good, or excellent (varies by author). Some researchers may ask the questions: 1) Is one parameter more "poor" than another and 2) How poor is "poor". Table 6 shows ICC values of 0 or less that correspond to CVs of 160.33% to 207.6%. In this case, the ICCs reported were very similar, but the CVs indicated a difference of approximately 47%. From this information, CVs help researchers to easily compare reliabilities of parameters (160.33% shows poor reliability, but not as poor as 207.6%). Researchers can also determine variation from test to test and directly apply the results (in this case, "poor" means that a given value can have a variation of up to 207.6%).

Another advantage of CVs comes with the interpretation of ICCs reliability (poor, fair, good, or excellent). Various authors apply different systems (Table 1), partly due to the implications of different reliability levels to their particular fields, but possibly partly due to the fluctuations of ICCs. With CVs, there are no fluctuations, and any researcher can easily see that 5% CV shows very high reliability, while 95% shows fairly low reliability. CV should be considered as one of the necessary reliability indexes to report.

CHAPTER 6. CONCLUSION

6.1 Summary and Applications

Maximum Voluntary Exertions for the anterior deltoid, middle deltoid, and trapezius showed excellent reliability. The methods for the three isometric tasks used are thus reliable, and can be used to obtain MVE for other tasks studying the shoulder muscles.

Using the Borg CR-10 scale to obtain Ratings of Perceived Discomfort resulted in excellent reliability for RPD slope and final rating. Thus, using RPD in a dynamic intermittent task to rate the shoulder should produce reliable results. Though more research is needed, this shows promise for applying RPD to other dynamic tasks. This would be especially useful and economical in industry. Sometimes obtaining fatigue readings through means such as EMG may be expensive and unpractical in a real world situation such as an automobile assembly line. RPD can provide a quick and reliable solution.

EMG intercepts, especially MnPF and MdPF, provided excellent reliability. On the other hand, EMG slopes were not very reliable parameters. MnPF and MdPF actual slope showed very low reliability. Normalizing these slopes did not seem to improve reliability, but taking the absolute value was more promising. Absolute slopes showed average reliability. Normalizing MnPF tended to increase reliability, while normalized MdPF tended to decrease reliability. MnPF normalized absolute slope was the most reliable Mn/MdPF slope parameter for the muscles studied.

RMS intercepts, though not as good Mn/MdPF, still provided average reliability. RMS actual slope showed very low reliability. Of all the slope parameters, MnPF absolute normalized slope and RMS absolute slope were the most reliable.

EMG parameters are useful for assessing muscle fatigue characteristics. Intercept and slope parameters show trends in EMG data that can be measured to make inferences about reliability. In this study, intercepts showed high reliability, and slopes showed poor to good reliability. This indicates that either EMG slopes are not reliable parameters of dynamic tasks involving the shoulder or more innovative procedures are necessary to represent trends in EMG data. In either case, more research in this area is required.

A summary of confidence intervals, applications, and limitations of the fatigue parameters studied is provided in Table 18. It should be noted that these parameters should be applied to tasks that are similar in nature to the present study and to the shoulder muscles studied (anterior deltoid, middle deltoid, and trapezius).

Table 18: Confidence Intervals, Applications and Limitations of Fatigue Parameters

Parameter	Confidence Intervals, Applications, and Limitations
Borg Final Rating	*95% C.I.*: ±1.74 *Application*: Borg Scale can be inexpensively applied to industrial tasks without much task interruption. Variation within 95% C.I should be considered normal, while final ratings above C.I. demands attention to industrial task as worker exhibits higher than normal discomfort ratings. *Limitation*: CV for higher final ratings (i.e. 7 to 10) is relatively high. This task was only studied for one hour and participants did not exhibit extreme discomfort; therefore, reliability of higher discomfort levels is undetermined.
Borg Slope	*95% C.I.*: ±.018 *Application*: Borg Scale can be inexpensively applied to industrial tasks without much task interruption. Variation within 95% C.I should be considered normal, while slopes above C.I. demands attention to industrial task as worker exhibits higher than normal discomfort ratings *Limitation*: Only useful for tasks with duration of one hour or less, and cannot be extrapolated to longer tasks without more research.
EMG Intercepts	*95% C.I.*: MnPF — Ant ±16.2Hz, Mid ±12.7Hz, and Trap ±17.8Hz MdPF — Ant ±19.3Hz, Mid ±11.3Hz, and Trap ±13.1Hz RMS — Ant ±.084, Mid ±.13, and Trap ±.088 *Application*: Variation within 95% C.I should be considered normal, while intercepts outside of C.I. may indicate inconsistent methods. *Limitation*: If slope is not reliable, then regression line of EMG data is not reliable, and regardless of the consistency of methods, the reliability of EMG data still is in question.
EMG Slopes	*95% C.I.*: N/A due to high reliability *Application*: Research has indicated that EMG slopes are reliable for some static tasks (Dedering et al., 2000), and may therefore be useful in industry. Taking the absolute value of slopes shows promise for reliability research and may be useful in studies involving other muscles for different dynamic tasks. *Limitation*: The slope EMG parameters in this study showed medium to low reliability, and thus, are not useful fatigue parameters for the muscles studied while performing an overhead work task.

Currently, the trend for reliability in biomechanics, as well as other areas, is to report ICC and SEM. Arguments may still be made over the usefulness of ICC or whether the alternative ICC equations should be used (Denegar et al., 1993). Regardless, given the advantages of CV, CV should be considered as a necessary reliability index.

6.2 Limitations

Sample size might be the most noticeable limitation in this study. Some researchers suggest anywhere from 8 participants (Hopkins, 2000) to 400 participants (Charter, 1999). There is an argument of an upper limit to the number of participants and a point wherein being too conservative and running more participants is a waste of time and funds, however, using only 10 participants for this study may be considered too liberal.

Another limitation, which can impact sample size and reliability, is the number of sessions. This study had two sessions (test and retest). Using more sessions (retests) can reduce the number of participants needed for an adequate reliability study. Donner et al. (1987) illustrates the relationship between number of sessions, sample size, and the detection of low or high reliability.

Another limitation is that the task was only conducted for one hour. None of the participants reached their endurance limit during this time. This might imply that none of the participants reached dramatically high levels of fatigue. If participants were to perform the task for a longer period of time, they may experience different fatigue characteristics. This may affect reliability parameters, especially EMG slope.

There are possibly other muscles, such as the supraspinatus, that are active in overhead tasks. Using EMG to measure these muscles and somehow combine these measurements with other active muscles may improve EMG slope reliability.

6.3 Future Research

There are several research possibilities that may be further explored. The most obvious is sample size and number of sessions. This study can be duplicated with more subjects and more sessions (minimum of three) to see if parameters show similar reliability.

Longer tests session (possibly until endurance limit) can be conducted as well. It may be more practical in this situation to adjust the task (increase mass or duty cycle) otherwise endurance times may become very lengthy. With longer session times, reliability can be calculated per minute or per cycle to determine at what point during the sessions do parameters decrease in reliability.

In order to improve reliability for EMG parameters, different approaches should be investigated. For example, Larsson (2003) not only reports reliability for Mn/MdPF intercepts/initial values, but also final values. For dynamic tasks, wherein workload for individual muscle (in a complex muscle group) may shift, understanding which muscles are prominent, and the distribution of the workload over the course of the task may help improve the reliability of EMG parameters.

REFERENCES

Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics*, **45**(14), 966-987.

Aaras, A., Veierod, M., Laren, S., Ortengren, R., and Ro, O. (1996). Reproducibility and stability of normalized EMG measurements on musculus trapezius. *Ergonomics*, **39**(2), 171-185.

Bartko, J., (1966). The intraclass correlation coefficient as a measure of reliability. *Psychol Rep*. **19**, 3–11.

Bjelle, A., Hagberg, M., and Michaellsson, G. (1981). Occupational and individual factors in acute shoulder and neck disorders among industrial workers. *British Journal of Industrial Medicine,* **38**, 356-363.

Bland, J. and Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, **8**, 307–310.

Björkstén, M.G., Boquist, B., Talback, M. and Edling, C. (2001). Reported neck and shoulder problems in female industrial workers: the importance of factors at work and at home. *International Journal of Industrial Ergonomics,* **27**, 159-170.

Borg, G. (1982). Psychophysical bases of perceived exertion. *Med Sci Sports Exercise*, **14**, 277-281.

Bureau of Labor Statistics (2001). U.S. Department of Labor.

Chen, Y. (2000). Changes in lifting dynamics after localized arm fatigue. *International Journal of Industrial Ergonomics*, **25**(6), 611-619.

Chaffin, D., Andersson, G., and Martin, B. (1999). Occupational Biomechanics 3rd Edition. New York: John Wiley & Sons, Inc., 43 – 45.

Charter, R. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology,* **21**(4), 559-566.

Cicchetti, D. (1999). Sample Size Requirements for Increasing the Precision of Reliability Estimates: Problems and Proposed Solutions. *Journal of Clinical and Experimental Neuropsychology*, **21**(4), 567-570.

De Luca, Carlo. (1997). The Use of Surface Electromyography in Biomechanics. *Journal of Applied Biomechanics,* **13**, 135-163.

Dedering, A., Nemeth, G., and Harms-Ringdahl, K. (1999). Correlation between electromyographic spectral changes and subjective assessments of lumbar muscle fatigue in subjects without pain from the lower back. *Clin Biomech*, **14**(2), 103-111.

Dedering, A., Roos af Hjelmaster, M., Elfving, B., Harms-Ringdahl, K., and Nemeth, G. (2000). Between-days reliability of subjective and objective assessments of back extensor muscle fatigue in subjects without lower-back pain. *Journal of Electromyography and Kinesiology*, **10**, 151-158.

Dempsey, P. (1998). A critical review of biomechanical, epidemiological, physiological and psychophysical criteria for designing manual materials handling tasks. *Ergonomics*, **41**(1), 73 – 88.

Denegar, C., and Ball, D. (1993). Assessing reliability and precision of measurement: An introduction to intraclass correlation and standard error of measurement. *Journal of Sport Rehabilitation*, **2**, 35-42.

Dimitrova, N.A., and Dimitrov G.V. (2003) Interpretation of EMG changes with fatigue: facts, pitfalls, and fallacies. *Journal of Electromyography and Kinesiology*, **13**(1), 13-36.

Donner, A., and Eliasziw, M. (1987). Sample size requirements for reliability studies. *Statistics in Medicine*, **6**(4), 441-448.

Ebenbichler, Gerold, R., Bonato, Paolo, Roy, Serge, H., Lehr, Stefan, Posch, Martin, Kollmitzer, Josef, Croce, D., and Ugo. (2002) Reliability of EMG time-frequency measures of fatigue during repetitive lifting. *Journal of American College of Sports Medicine*, **34**(8), 1316-1323.

Elfving, B., Liljequist, D., Mattsson, E., and Nemeth, G. (2002). Influence of interelectrode distance and force level on the spectral parameters of surface electromyographic recordings from the lumbar muscles. *Journal of Electromyography and Kinesiology*, **12**, 295-304.

Elfving, B., Nemeth, G., Arvidsson, I., and Lamontagne, M. (1999). Reliability of EMG spectral parameters in repeated measurements of back muscle fatigue. *Journal of Electromyography and Kinesiology*, **9**, 235-243.

Fagarasanu, M., and Kumar, S. (2002). Measurement instruments and data collection: a consideration of constructs and biases in ergonomics research. *International Journal of Industrial Ergonomics*, **30**(6), 355-369

Falla, D., Dall'Alba, P., Rainoldi, A., Merletii, R., and Jull, G. (2002). Repeatability of surface EMG variables in the sternocleidomastiod and anterior scalene muscles. *European Journal of Applied Physiology.*

Flatow E., Soslowsky, L., and Ticker, J. (1994). Excursion of the rotator cuff under the acromion: patterns and subacromial contact. *American Journal of Sports Medicine,* **22,** 779-788.

Fleiss, J. (1986). The Design and Analysis of Clinical Experiments. New York: John Wiley & Sons.

Gerdle, B., Larssona, B., and Karlssonc, S. (2000). Criterion validation of surface EMG variables as fatigue indicators using a peak torque – a study of repetitive maximum isokinetic knee extensions. *Journal of Electromyography and Kinesiology*, **10**(4), 225-232

Green, H.G. (1996). Mechanisms of muscle fatigue in intense exercise. *Journal of Sports Science*, **15**, 247-256.

Gorman, R., McKenzie, D., Gandevia, S. (1999). Task failure, breathing discomfort and CO2 accumulation without fatigue during inspiratory resistive loading in humans. *Respiration Physiology*, 115, 273–286.

Hopkins, W. (2000). Measures of reliability in sports medicine and science. *Sports Med*. **30**(1), 1-15.

Hui, L., Ng, G., Yeung, S., and Hui-Chan, C. (2001). Evaluation of physiological work demands and low back neuromuscular fatigue on nurses working in geriatric wards. *Applied Ergonomics*, **32**(5), 479-483.

Ilg, A., Bankes, M., and Emery, R. (2001). The intra- and inter-oberservor validity of the Samilson and Preito grading system of glenohumeral anthropathy. *Knee Surg, Sports Traumatol, Arthrosc*, **9**, 187-190.

Kim, J., Chung, M., and Park, J. (2003). Measurement of physical work capacity during arm and shoulder lifting at various shoulder flexion and ad/abduction angles. *Human Factors and Ergonomics in Manufacturing*, **13**(2), 153-163.

Keating, J. (1998). Unreliable inferences from reliable measurements. *The Australian Journal of Physiotherapy*. **44**, 5-10.

Kollmitzer, J., Ebenbichler, G., and Kopf, A. (1999). Reliability of surface electromyographic measurements. *Clinical Neurophysiology*, **110**(4), 725-734.

Koumantakis, G., and Oldham, J. (2002). Between-days reliability of electromyographic measures of paraspinal muscle fatigue at 40, 50, and 60% levels of maximum voluntary contractile force. *Clinical Rehabilitation*, **16**, 761-771.

Kumar, S., Narayan, Y., and Bjornsdottir, S. (1999). Comparison of the sensitivity of three psychophysical techniques to three manual materials handling task variables. *Ergonomics*, **42**(1), 61-73.

Landis, J., and Kock, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159-174.

Lariviere, C., Gagnon, D., and Loisel, P. (2000) The comparison of trunk muscles EMG activation between subjects with and without chronic low back pain during flexion-extension and lateral bending tasks. *Journal of Electromyography and Kinesiology*, **10**, 79-91.

Lariviere, C., Arsenault, B., Gravelb, D., Gagnonc, D. and Loiseld, P. (2002) Evaluation of measurement strategies to increase the reliability of EMG indices to assess back muscle fatigue and recovery. *Journal of Electromyography and Kinesiology*, **12**(2), 91-102.

Larsson, B., Karlssona, S., Erikssona M., and Gerdle B. (2003). Test–retest reliability of EMG and peak torque during repetitive maximum concentric knee extensions. *Journal of Electromyography and Kinesiology,* **13**(3), 281-287

Larsson, B., Mansson, B., Karlber, C., Syvertsson, P., Elert, J., and Gerdle, B. (1999). Reproducibility of surface EMG variables and peak torque during three sets of ten dynamic contractions. *Journal of Electromyography and Kinesiology*, **9**, 351-357.

Lohr, J. and Uhthoff, H. (1990). The microvascular pattern of the supraspinatus tendon. *Clinical Orthopedics and Related Research,* **254,** 35-38

Mador, M., Rodis, A., and Magalang, U. (1995). Reproducibility of Borg scale measurements of dyspnea during exercise in patients with COPD. *Chest*, **107**(6), 1590-1598.

Mannion, A., and Dolan, P. (1994). Electromyographic median frequency changes during isometric contraction of the back extensors to fatigue. *Spine*,**19**(11), 1223-1229.

Nargol, A., Jones, A., Kelly P., and Greenough C. (1999). Factors in reproducibly of electromyographic power spectrum analysis of lumbar paraspinal muscle fatigue. *Spine*. 1999, **24**, 883-888.

National Institute of Occupational Safety and Health. (1997). Musculoskeletal Disorders (MSDs) and Workplace Factors.

Nussbaum, M. (2001). Static and dynamic myolectric measures of shoulder muscle fatigue during intermittent dynamic exertions of low to moderate intensity. *Eur J. Appl. Phsyiol,* **85**, 299-309

Nussbaum, M., Clark, L., Lanza, M., and Rice., K. (2001). Fatigue and Endurance Limits During Intermittent Overhead Work. *AIHAJ*, **62**, 446-456.

Occupational Safety and Health Administration. (1999). Report of the Small Business Advocacy Review Panel on the Draft Proposed Ergonomics Program Rule.

Peach, J., Gunning, J., and McGill, S. (1998). Reliability of spectral EMG parameters of healthy back extensors during submaximum isometric fatiguing contractions and recovery. *Journal of Electromyography and Kinesiology,* **8**(6), 403-410.

Perkins, D., Wyatt, and R., Bartko, J. (2000). Penny-wise and Pount-foolish: The Impact of Measurement Error on Sample Size Requirements in Clinical Trials. *Biol Psychiatry*, **47**, 762-766.

Radwin, R., Marras, W., and Lavender, S. (2002). Biomechanical aspects of work-related musculoskeletal disorders. *Theoretical Issues in Ergonomics*, **2**(2), 153-217.

Rainoldi, A., Bullock-Saxton, J., Cavaretta, F., and Hogan, N. (2001). Repeatability of maximal voluntary force and of surface EMG variables during voluntary isometric contraction of quadriceps muscles in healthy subjects. *Journal of Electromyography and Kinesiology*, **11**, 425-438.

Rainoldi, A., Galardi, G., Maderna, L., Comi, G., Lo Conte, L., and Merletti, R. (1999). Repeatability of surface EMG variables during voluntary isometric contractions of the biceps brachii muscle. *Journal of Electromyography and Kinesiology*, **9**, 106-119.

Rew, L., Koniak-Griffin, D., Lewis, M., Miles, M., O'Sullivan, A., and American Academy of Nursing's Expert Panel on Adolescent Health. (2000). Secondary data analysis: New perspective for adolescent research. *Nursing Outlook*, **48**(5), 223-229.

Roy, S., De Luca, C., and Casavant, D. (1989). Lumbar muscle fatigue and chronic lower back pain. *Spine*, **14**, 992–1001.

Salerno, D., Franzblau, A., Armstrong, T., Werner, R., and Becker, M. (2001). Test-retest reliability of the upper extremity questionnaire among keyboard operators. *American Journal of Industrial Medicine*, **40**, 655-666.

Shen, W., and Parsons, K. (1997). Validity and reliability of rating scales for seated pressure discomfort. *International Journal of Industrial Ergonomics*, **20**, 441-461.

Shrout, P., and Fleiss, J. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, **86**(2), 420-428.

Shrout, P. (1998). Measurement reliability and agreement in psychiatry. *Statistical methods in Medical Research*, **7**, 301-317.

Sleivert, G., and Wenger, H. (1995) Reliability of measuring isometric and isokinetic peak torque, rate of torque development, integrated electromyography, and tibial nerve conduction velocity. *Arch Phys Med Rehabil*, **75**,1315–1321.

Sood, S. Hager, K. and Nussbaum, M. (2002) The effects of differing overhead heights on shoulder fatigue during a repetitive intermittent task. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting.* Baltimore, MD. pp 1081-1085.

Stokdijk, M., Beigstraatan, M., Ormel, W., Boer, Y., Veeger, H., and Rozing, P. (2000) Determining the optimal flexion-extension axis of the elbow in vivo – a study of interobserver and intraobserver reliability. *Journal of Biomechanics*, **33**, 1139-1145.

Viitasalo, J., and Komi, P. (1975) Signal characteristics of EMG with special reference to reproducibility of measurements. *Acta physiol. Scand.*, **93**, 531-539.

Wikstrom, B. (1993). Effects from twisted postures and whole-body vibration during driving. *International Journal of Industrial Ergonomics*, **12**, 61-75.

Wos, H., Marez, T., Noworoi C., and Borg G. (1988). The reliability of self-ratings based on Borg's Scale for hand-arm vibrations of short duration. *International Journal of Industrial Ergonomics*, **2**, 151-156.

APPENDIX A: INFORMED CONSENT PACKAGE

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING (ISE)

Informed Consent for Participants of Investigative Projects

Title of Project: "Recommended Limits for Overhead Work Tasks"

Principal Investigators: Dr. M. A. Nussbaum, Associate Professor, ISE
Kristopher Hager, Graduate Research Assistant, ISE
Deepti Sood, Graduate Research Assistant, ISE

I. THE PURPOSE OF THIS RESEARCH

You are invited to participate in a study to determine limits for upper-extremity overhead work related to automotive assembly tasks. To obtain this information, two experiments are to be conducted. The first is designed to determine the influence of task height. The second experiment will involve simulation of overhead work in a variety of conditions. It is anticipated that there will be approximately 6 participants for the first experiment, and 46 participants for the second experiment (6 of which are pilot studies).

II. PROCEDURES

The procedures used in this study are as follows.

1) You will have electrodes placed on several muscles, which move the shoulder. These electrodes are used to collect information from the muscles, which can indicate fatigue levels. The procedure for each electrode involves cleansing a small patch of skin (approximately the size of two quarters) over the muscle area. The electrodes are then placed on the skin and remain in place with an adhesive.
2) The investigator will demonstrate the data collection procedures, which involve performing overhead work tasks at various heights, or performing overhead work tasks at various work-cycle durations and exertion levels.
3) You will conduct simulated overhead work cycles as demonstrated by the investigator with rest periods after each exertion.
4) For the first experiment, each participant will perform simulated overhead work for a maximum of one hour at three different heights.
5) For the second experiment, each of participants will perform simulated overhead work for 10 minutes in each of 8 different task conditions, with 5 minutes rest between each condition.

The total estimated time of participation is 2 hours (including rest periods) for the first experiment, and 3 hours (including rest periods) for the second experiment.

III. RISKS AND BENEFITS OF THIS RESEARCH

Your participation in this study will provide information that will be used to develop design guidelines for overhead work. It is the objective of this study to contribute design information for improving worker safety, comfort, and productivity.

The primary focus of this study is to measure muscle fatigue. Therefore, you may experience some discomfort related to extended use of some muscles. The muscle fatigue will occur due to use over a long period of time with regular breaks, and not due to generation of large forces. In addition, an investigator will continuously monitor your condition to minimize any opportunity of strain.

There is minimal risk involved in this study.

IV. EXTENT OF ANONYMITY AND CONFIDENTIALITY

It is the intent of the investigators of this project to report the findings of this study. The information you provide will have your name removed and only a subject number will identify you during analysis and any written reports of the evaluation.

V. COMPENSATION

If you decide to participate in this study, you will be paid $10.00 per hour for the time you participate. The evaluation is expected to last 2-3 hours depending on the experiment. You will be paid at the conclusion of the testing session.

VI. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason without penalty. If you choose to withdraw during the study, you will be compensated for the portion of the testing which has been completed.

VII. APPROVAL OF THIS RESEARCH

This research project has been approved, as required, by the Institutional Review Board for projects involving human participants at Virginia Polytechnic Institute and State University, and by the Grado Department of Industrial Engineering.

VIII. PARTICPANT RESPONSIBILITIES

I know of no reason why I cannot participate in this study. I have the following responsibilities:

- To notify the investigator at any time about a desire to discontinue participation.

- To notify the investigator of any medical conditions which may be negatively influenced by extended muscular exertion. This may include heart disease, conditions influenced by blood sugar levels, or any other medical problems that may interfere with results or increase the risk of injury or illness.

____

Signature of Participant

IX. PARTICIPANT'S PERMISSION

Before you sign the signature page of this form, please make sure that you understand, to your complete satisfaction, the nature of the study and your rights as a participant. If you have any questions, please ask the investigator at this time. Then, if you decide to participate, please sign your name above and on the following page (please repeat for your copy).

Signature Page

I have read a description of this study and understand the nature of the research and my rights as a participant. I hereby consent to participate, with the understanding that I may discontinue participation at any time if I choose to do so.

Signature ___________________________

Printed Name ________________________

Date ____________________________

The research team for this experiment includes Dr. M. A. Nussbaum, Assistant Professor, Kristopher Hager, Graduate Research Assistant, and Deepti Sood, Graduate Research Assistant. Research team members may be contacted at the following address and phone number:

Grado Department of Industrial and Systems Engineering Department
250 New Engineering Building
Virginia Tech
Blacksburg, VA 24061
(540) 231-6053

In addition, if you have detailed questions regarding your rights as a participant in University research, you may contact the following individual:

Dr. David Moore
Chair, Institutional Review Board
CVM Phase II (Pathobiology)
Virginia Tech
Blacksburg, VA 24061
(540) 231-4991

Description of Recommended Limits for Overhead Work Tasks

Experimental Protocol for IRB Consideration

Submitted by:

Dr. Maury A. Nussbaum, Industrial Ergonomics Laboratory,
Assistant Professor
Phone: 231-6053, Email: nussbaum@vt.edu
(Short biography attached)

Ergonomic design tries to ensure that workers have sufficient capacity to perform required tasks and that completion of these tasks does not impose increased risk of musculoskeletal disorders (MSDs). This study will yield a practical design tool for application to automotive assembly tasks that involve upper extremity overhead work. The end product will be a set of data, in the form of tables, and statistical distributions, describing acceptable task durations for overhead work. This study will require the participants to be present for one to two sessions will last approximately two to three hours depending on the experiment. This includes rest periods during each session.

Several types of data will be collected from the subjects in the following manner.

1) Electromyography (EMG) data will be collected from major muscles recruited during the task. Standard non-invasive EMG procedures will be employed utilizing surface electrodes.
2) Verbal responses will be elicited from the participant through interview questions regarding the participant's perceived state of fatigue.
3) The participant will be prompted to rate their discomfort levels throughout the sessions. The Borg CR-10 scale, which includes verbal anchors for discomfort levels, will be visible to participants throughout the sessions.

This information will be gathered while the participant performs a simulated overhead working task. In the first experiment, participants will perform a single simulated overhead work task at various heights ranging from the participant's height to the participant's maximum overhead extended reach. In the second experiment, each participant will be tested under eight different conditions, which will vary work-cycle duration and vary exertion levels.

APPENDIX B: DEMOGRAPHIC AND ANTRHOPOMETRIC FORMS

Subject Data

Participant Number ________________ **Date** ________________

Starting Time ___________ **Age** ________________

Participant Name ____

First Middle Last

Address

Home ____

Street Apt #

____

City State Zip Code

Office ____

Street Apt #

____

City State Zip Code

Contact Number Home ___________________________________

Office ___________________________________

Email ID ____

Gender
- Male
- Female

Dominant Hand
- Right
- Left

Ethnicity
- Caucasian
- African Americans
- Asians/Pacific Islanders
- Native Americans
- Hispanics
- Other

Form A Anthropometric and Workstation Data

Anthropometrics

Weight _______________ kg

Stature _______________cm

Shoulder Height _______________ cm

Upper Arm Length ______________ cm

Lower Arm Length _______________ cm

A = Arm fully raised (holding tool) _________ cm

B = Arm at 90^o (holding tool) ____________ cm

Workstation Height

H1’ = B = ________ cm

H2’ = (A-B)*0.40 = ________ cm

H3’ = (A-B)*0.80 = ________ cm

Task MVC Height

H1” = B – T = ________ cm

H2” = H2’– T = ________ cm

H3” = H3’– T = ________ cm

H1 = B + 42.7 = ______________________________ = ____________cm

H2 = H2’ + 42.7 = _____________________________= _____________cm

H3 = H3’ + 42.7 = _____________________________= _____________cm

PLACE ELECTRODES!!

Form B Demographics and Musculoskeletal Data

Demographics

1. Present Occupation (Part/Full time) ____
2. How many hours per week? ____
3. Previous Occupation (Part/Full time) ____
4. Description of Manual Work ____
____
5. How long have you done Manual Work Occupation? ________________________________
6. Have you had a significant injury? ____
7. If yes, which body parts were affected by the injury? ________________________________
8. How would you describe your general fitness level?

• Minimal • Moderate • Average • Above Average • Maximal

Musculoskeletal Trouble

Have you had Pain, Ache, Discomfort, Injuries in:	In the past 12 months		In the last 7 days	
	When did it occur	Duration It lasted	When did it occur	Duration It lasted
Neck				
Shoulders				
Elbows / Wrist / Hands				
Upper Back / Lower Back				
Knees / Legs				
Hips / Thighs				
Knees / Ankles / Feet				

Any Other Comments: ____
____

APPENDIX C: EMG, MVE, AND RPD FORMS

Form D Experiment Data Sheet

A. Electrodes (Impedance)

AD ___________ (< 10k Ohm?)

MD___________ (< 10k Ohm?)

TR ___________ (< 10k Ohm?)

B. Electrode Distance

Electrode distance (Acromion-b/w Markers)___________ cm

Electrode distance (Acromion-b/w Markers)___________ cm

Electrode distance (Acromion-b/w Markers)___________ cm

C. Recording (EMG Gain)

AD ________ MD ________ TR ________

D. MVE'S per Trial

Trials	Anterior Deltoid	Middle Deltoid	Trapezius	Task
T1				
T2				
T3				
T4				
T5				
T6				
T7				

Comments/Unusual Circumstances ____

____

____

Form E Borg Scale

Height #							
Cycle No.	**Shoulder (Borg Scale)**	**Cycle No.**	**Psycho-physical Prediction**	**Cycle No.**	**Back (Low, Med, Hi)**	**Cycle No.**	**Neck (Low, Med, Hi)**
2		3		4		5	
7		8		9		10	
12		13		14		15	
17		18		19		20	
22		23		24		25	
27		28		29		30	
32		33		34		35	
37		38		39		40	
42		43		44		45	
47		48		49		50	
52		53		54		55	
57		58		59		60	
62		63		64		65	
67		68		69		70	
72		73		74		75	

T. Number of Full Trials Completed: ______

Stopped During Trial Number: ______

VITA

Kristopher M. R. Hager

Kristopher Hager completed a B.S. in Industrial and Systems Engineering at Virginia Tech in 2001, and a M.S. in Industrial and Systems Engineering with a focus on Biomechanics at Virginia Tech, graduating in 2003. While pursuing his M.S. degree, he worked as research assistant in Virginia Tech's Industrial Ergonomics Laboratory, and as a teaching assistant. Kristopher was president of the Human Factors and Ergonomics Society (HFES) Virginia Tech student chapter, wherein he helped guide the chapter to win the HFES Outstanding Student Chapter of the Year award in 2002. In 2003, Kristopher began work as a business and ergonomic analyst in Evans, Inc. in Vienna, Virginia.